

AUG 08 2003

TECH CENTER 1600/2900

[0001] The present application is the subject of Provisional Application Serial No. 60/208,650 filed June 2, 2000 entitled ALGORITHMIC DETERMINATION OF CONNECTRONS FOR THE HIGH LEVEL REGULATION OF GENE EXPRESSION.

[0002] RNA introduced into a cell by a virus is now known to trigger a cellular defense mechanism known as post-transcriptional gene silencing (PTGS). If the viral RNA sequence matches a sequence within the cell's genome the associated genes are turned off or silenced. This phenomenon is also called 'RNA interference' or RNAi. A single-stranded RNA can interact with another single-stranded RNA (known as antisense RNA). The single-stranded RNA can also form a triple-stranded complex with double-stranded DNA. This triple-stranded complex is known as a Hoogsteen helix. This patent application shows how two specific adjacent RNA single-stranded sequences (called C1 and C2 - for Control Sequence 1 and Control Sequence 2) interact with two distant double-stranded DNA sequences (called T1 and T2 - for Target Sequence 1 and Target Sequence 2) to form a tetradic relationship which is called a "connectron". The two distant DNA double-stranded sequences (T1 and T2) must be on the same chromosome in a genome and they must be between about 1kb and 105kb of each other. The adjacent single-stranded RNA sequences (C1/C2) can be on the same or different chromosome as the T1 and T2

sequences. The C1 sequence is identical to the T1 sequence and the C2 sequence is identical to the T2 sequence. The connectron acts to stabilize the double-stranded DNA by allowing 30nm chromatin particles to form. Genes that lie between the T1 and T2 sequences when wrapped up in 30nm chromatin particles are not open to promotion and expression. The connectron (i.e. the tetradic relationship between the T1-T2 sequences and C1/C2 sequences) provides a general explanation for PTGS. A connectron can be implemented by RNA sequences, PNA (Peptide Nucleic Acid) sequences or by a zinc-finger DNA Binding Protein (DBP) specific to the T1 and T2 sequences.

[0003] Characteristically the adjacent C1/C2 sequences lie in the 3'UTR of a gene. The T1 and T2 sequences do not lie within the translated region of any gene. These sequences "surround" one or more genes. There are, however, T1 and T2 sequence pairs that surround one or more C1/C2 sequences that are not 3'UTR to any gene. These are called "geneless connectrons". There may be promoter sequences that cause the transcription of these 3'UTR sequences.

[0004] A computer-based algorithm that is similar to the algorithm used in the US Patent 6,205,404 has been developed to determine the connectron structure of any genome. This algorithm determines the existence of all the connectrons in the genomic DNA. Connectrons exist in prokaryotes, archaea, single-celled eukaryotes, multi-celled eukaryotes, plants and higher animals. Connectron relationships exist between prokaryotes and their plasmids. The geneless connectrons provide a possible mechanism for forming a hierarchy of gene expression control that will produce an understanding of cell differentiation and tissue development.

[0005] Each connectron is a unique tetrad of sequences. Each connectron changes the expression of the genes between the T1 and T2 sequences. The C1 sequence (which is equivalent to the T1 sequence) and the C2 sequence (which is equivalent to the T2 sequence) are determined by the invention described in this patent application. In general, the tetrad of connectron sequences can be patented because the structure of matter is known and the function of specific gene expression modulation is also known. Gene expression modification can be produced by introducing antisense RNA or PNA to interact C1/C2 RNA sequences or zinc-finger DBPs to interact with the T1 and T2 sequences. Using connectrons it will be possible to modify cellular and tissue behavior in a very general manner.

[0006] Examples will be given from different genomes to illustrate that the connectron is a perfectly general and universal concept.

Definitions

[0007] Double stranded DNA - Watson and Crick showed in 1953 that DNA naturally forms a double-stranded helix. A typical double stranded sequence is

[0008] 5'-TAGAGGAGTACCAC-3'

[0009] 3'-ATCTCCTCATGGTG-5'

[00010] Hydrogen Bond - The force between a hydrogen atom and another heavier atom such as Oxygen (O), Nitrogen (N), Phosphorus (P), or Sulfur (S).

[00011] Positive strand - The positive strand is normally represented 5' to 3' running left to right as in

[00012] 5'-TAGAGGAGTACCAC-3'

[00013] Negative strand - The negative strand is normally represented 5' to 3' running right to left as in

[00014] 3'-ATCTCCTCATGGTG-5'

[00015] Single stranded RNA - Either the positive or the negative strand of the double-stranded DNA can be transcribed by the polymerase. In RNA U replaces T.

[00016] RNA of positive strand sequence 5'-AGAGGAGUACCAC-3'

[00017] RNA of negative strand sequence 5'-GUGGUACUCCUCUA-3'

[00018] Antisense RNA - The antisense strand of any RNA sequence is the compliment sequence

[00019] RNA sequence 5'-UAGAGGAGUACCAC-3'

[00020] Antisense RNA sequence 3'-AUCUCCUCAUGGUG-5'

[00021] Triple Strand Helix - The RNA sequence of a RNA/DNA triple-strand complex is the same as the positive strand of the DNA

[00022] DNA positive strand 5'-TAGAGGAGTACCAC-3'

[00023] DNA negative strand 3'-ATCTCCTCATGGTG-5'

[00024] RNA strand 5'-UAGAGGAGUACCAC-3'

[00025] Promoter - Any region of DNA, that binds proteins which engage the polymerase transcription mechanism.

[00026] TATA Box - A region near the 3' end of a promoter with the sequence TATA.

[00027] mRNA - The RNA produced from the DNA by the polymerase as a result of transcription

[00028] Start of transcription - The 3' end of a promoter where the polymerase mechanism begins to transcribe DNA into mRNA.

[00029] Exon - Any region of mRNA which is used to code for proteins

[00030] Intron - Any region of mRNA lying between two exons which is not used to code for proteins. The introns are edited out of the initial RNA transcript to form the mature mRNA.

[00031] 3' UTR - The untranslated 3' end of an mRNA is beyond the end of the last exon. A stop codon in the mRNA causes the ribosome to stop the translation of mRNA into protein.

[00032] End of translation - The 3' end of the 3'-most exon.

[00033] Translated region - Any collection of exons and introns.

[00034] Gene - Any DNA region that codes for a protein. Introns do not occur in prokaryotic genes and they sometime fail to occur in eukaryotic genes. A typical model of a gene is

```
[00035]  |<-----Promoter----->|
          |<-TATA Box->|
          |<-Beginning of Translation
          |<-----Translated Region----->|
                        End of Translation-
          |<-Exon->|<-Intron->|<-Exon->|<-Intron->|<-Exon->|<-3'UTR->|
```

```

+ strand -----
- strand -----
|<-----Gene----->|

```

[00036] Positive strand gene - Any gene in which the features run 5' to 3' on the positive strand

[00037] Negative strand gene - Any gene in which the features run 5' to 3' on the negative strand

[00038] C1 sequence - Any positive or negative strand DNA sequence of 20 bases or more.

[00039] The C2 sequence must occur in the same chromosome as the C1 sequence.

[00040] C2 sequence - Any positive or negative strand DNA sequence of 20 bases or more.

[00041] The C1 sequence must occur in the same chromosome as the C2 sequence.

[00042] C1/C2 - Any positive or negative strand DNA sequence of 40 or more bases such that the C1 sequence is adjacent to the C2 sequence

[00043] T1 sequence - Any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence. The T1 and T2 sequences must be between about 1kb and 105kb apart.

[00044] T2 sequence - Any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence. The T2 and T1 sequences must be between about 1kb and 105kb apart.

[00045] Last exon gap or Gap-Distance - The number of bases between the end of transcription and the beginning of the C1/C2 sequence. In prokaryotes and single-celled eukaryotes this gap can range from no bases to 500 bases. In multi-celled eukaryotes the gap can be as large as 10,000 bases.

[00046] Poly-adenylation signal - A number of Adenosine (A) bases are added to the mRNA at the end of the 3'UTR.

[00047] Possible Connectron - Any set of T1, T2 and C1/C2 sequences such that the C1 sequence is identical to the T1 sequence and the C2 sequence is identical to the T2 sequence. The promoter of some gene causes the mRNA of the gene to be expressed. The mRNA is edited to eliminate the introns. The whole mRNA including the 3'UTR can move about in the cell or the nucleus of the cell. The C1/C2 RNA that is part of the 3'UTR moves to the T1 and T2 DNA sequences. A triple-stranded complex of the DNA and the RNA forms such that the C1 sequence forms hydrogen bonds with the T1 sequence and the C2 sequence forms hydrogen bonds with the T2 sequence. Because the C1 sequence is adjacent to the C2 sequence, the T1 sequence is brought physically close to the T2 sequence. This produces a loop of between about 1kb and 105kb in the DNA. Histone proteins reduce the length of the DNA by binding 200 bases. Histone/DNA complexes form six-fold symmetry chromatin assemblies. The diameter of the chromatin assemblies is approximately 30nm.

[00048] Real Connectron - Any Possible Connectron which is within the Gap-Distance of some gene

[00049] Homologous connectron - The T1 sequence and the T2 sequence are on the same chromosome as the C1/C2 sequence

[00050] Heterologous connectron - The T1 sequence and the T2 sequence are on a chromosome different from chromosome of the C1/C2 sequence

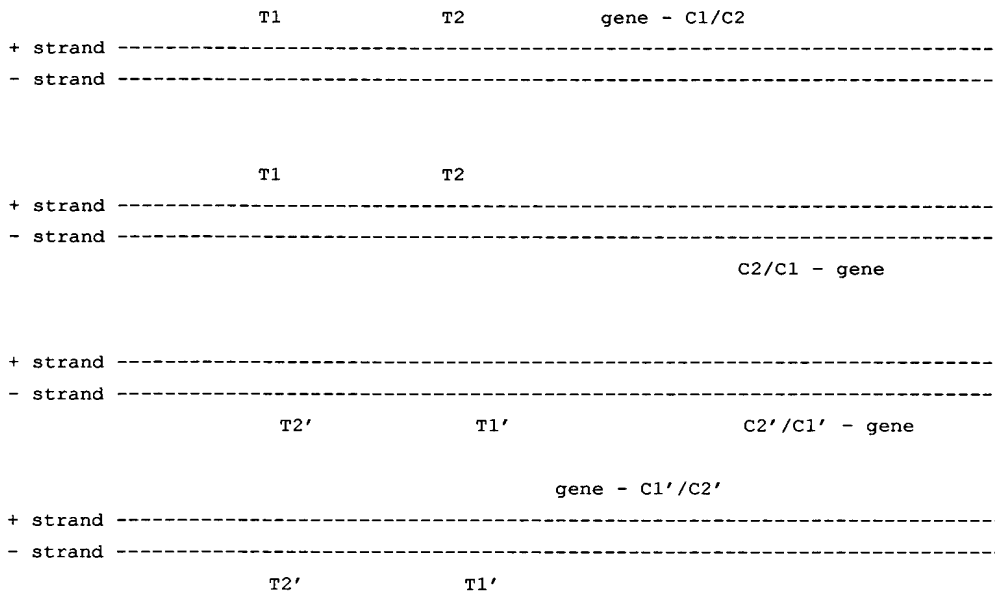
[00051] Permanent connectron - Any C1/C2 sequence, which is 3' UTR to some gene that is not surrounded by any T1 and T2 sequence pairs

[00052] Transient connectron - Any C1/C2 sequence, which is 3' UTR to some gene that is surrounded by one or more T1 and T2 sequence pairs

[00053] Self-limiting connectron - Any C1/C2 sequence which is 3'UTR to some gene that is surrounded by the T1 and T2 sequences such that C1=T1 and C2=T2

[00054] Geneless connectron - Any C1/C2 sequence which is not 3'UTR to some gene but is surrounded by some T1 and T2. A promoter may lie 5' to the C1/C2 sequence.

[00055] Bidirectionality of Connectron Excitation - A C1/C2 short loop on one strand selects a T1-T2 long loop pair on the same or the opposite strand. The C1/C2 short loop has a complementary C1'/C2' sequence on the opposite strand. Similarly the T1-T2 long loop pair has a complementary long loop pair T1'-T2'. Wherever a C1/C2, T1-T2 tetrad exists there is a complementary C1'/C2', T1'-T2' tetrad. The C1/C2 short loop can be transcribed as a 3'UTR to a gene on the same strand. The C1'/C2' short loop which is on the strand opposite to the C1/C2 short loop can also be transcribed as a 3'UTR to a gene on the same strand. There are four possible models of action



[00056] Of course, the short loops and the long loops do not have to be on the same chromosome.

[00057] Hierarchy of connectron action - When a C1/C2 is expressed it forms a T1-T2 loop by forming a connectron. The C1/C2 sequence does not have to be on the same chromosome as the T1 and T2 sequences. This provides a way of causing interaction between chromosomes. When the T1-T2 loop forms, any genes in that loop region which had been expressing C1/C2 sequences in their 3'UTRs, now cease expressing the C1/C2 sequences. The connectrons formed by these C1/C2 sequences will cease to exist after some time thus opening up the genes inside the respective T1-T2 loops to expression. The hierarchy of connectron action alternates between repression and expression. The connectron hierarchies can be of any depth.

[00058] One-to-Many connectron action - One C1/C2 sequence can form connectrons in many different places on many different

chromosomes. The only requirement is that $C1=T1$ and $C2=T2$. This makes it possible for one expression event to control the expression of many genes on different chromosomes.

[00059] Many-to-One connectron action - $C1/C2$ s that come from many different places on many different chromosomes can form a connectron for a specific $T1-T2$ sequence pair. The only requirement is that $C1=T1$ and $C2=T2$. This makes it possible for many different expression events to control the expression of one set of genes on a particular chromosome.

[00060] Many-to-Many connectron action - The arrangement of $C1/C2$ s and $T1-T2$ s across chromosomes can form a complex web of gene expression control relationships.

[00061] Percentage of the Genome Regulated by Connectrons - Since the connectrons for a sequenced genome can be calculated, the percentage of the genome that is open to connectron regulation can be known.

[00062] Emergent Property - The network of connectrons in any genome emerges from a knowledge of the complete DNA sequence of the genome. Because both the $C1/C2$ sequences and the $T1-T2$ sequences can be any place in the genome, the whole genomic sequence must be known before all the connectrons can be determined.

[00063] Paradigm Shift - For the past fifty years since the discovery by Watson and Crick of the double-helical nature of DNA, the reigning paradigm for scientific discovery has been the study of one gene and its effects on the behavior of a cell. The advent of genomic sequencing and this invention of connectrons that emerge from the whole genome will produce a shift in the way scientists view biological systems and the

way they formulate and execute experiments. The many-to-many relationships between the connectrons means that there are many ways in which the expression of a set of genes can be modulated. The multiplicity of control pathways means produces a system stability that makes it possible for biological systems to be stable for long periods of evolutionary time. The thinking that goes into formulating scientific experiments will have to change to accommodate the changes in understanding that will be induced by the application and extension of this patent application.

[00064] Hierarchy of DNA Structuring - The DNA of a cell's genome is structured in a hierarchy of six levels. Figures 1, 2 and 3 have been adapted from The Molecular Biology of the Cell by Alberts, Bray, Lewis, Raff, Roberts and Watson [third edition pages 354, 345 and 348]. As shown in figure 1, the double stranded DNA is level 1. The double-stranded DNA is wrapped around histone proteins to form a chromatin particle that is level 2 of the hierarchy. Level 2 is described as "beads-on-a-string" in figure 1. The chromatin particles are packed in a six-fold symmetry as shown in figure 2a and figure 2b. These six-fold assemblies have a diameter of 30 nm. Each 30 nm assembly contains from 18 (i.e. $6 * 3$) to 30 (i.e. $6 * 5$) chromatin particles. The 30 nm assemblies aggregate into large loops which range in length from 5,000 bases to 100,000 bases of DNA. The size of these large loops as shown in figure 1 is approximately 300 nm. These large loops constitute level 4 of the structuring hierarchy. As shown in figure 1, level 5 of the DNA structuring hierarchy many large loops are condensed to form a structure which is approximately 700 nm in diameter. The complete chromosome that constitutes level 6 of the hierarchy is composed of two very long sections of level 5 DNA.

[00065] Model of Chromatin Structure - The level 4 structure of DNA as shown in figure 1 ranges in length from 5,000 to 105,000 bases of DNA. Figure 3 shows that proteins are thought to connect portions of the long loops formed by the 30 nm particles to form a chromosome axis. These condensed long loops are described as chromomeres in The Molecular Biology of the Cell.

Prior Art

[00066] The chromomere model of DNA structuring was presented by N. A Resnik, et al.[1] and is based on electron microscopic data. There are more recent papers studying a variety of genomes with electron microscopy but no equivalent study of chromomeres has been done on a fully sequenced genome.

[00067] A recent News Feature in Nature by T. Gura [2] described the discovery of post-transcriptional gene silencing in which viral RNA interacts with the transcribed RNA of the cell to silence the expression of genes. This article describes experiments in *C. elegans* and *D. megalomaster* in which RNA that is complementary to mRNA introduced into a cell. This "antisense" RNA has the effect of turning off the expression of one or more genes. The introduced complementary RNA produces an "RNA interference" called RNAi.

[00068] Thomas Werner and his colleagues at Genomatix in Munich, Germany have developed an approach to understanding what they call "Matrix Attachment Region" (MAR). Figure 5 shows their interpretation of the structure of DNA surrounding a gene. The following description of the MAR is copied from the Genomatix web site

[00069] "Matrix Attachment Regions (MARs) MARs are sequence regions that are responsible for the attachment of genomic DNA to the nuclear matrix or scaffold. Transcription absolutely requires anchorage of genomic DNA to the nuclear matrix.

Functional features of MARs:

Anchoring of regulatory elements like promoters and enhancers to the nuclear matrix.

Ensuring long term activity of promoters and enhancers in chromatin.

Insulation, rendering a functional domain insensitive to position effects.

[00070] Genomatix is conducting a research project to define and detect MARs by computer-analysis."

Brief Description of the Objects of the Invention

[00071] An object of the invention is to provide a method of identifying DNA sequences that control the expression of different collections of genes in a genome comprising, detecting selected DNA sequences adjacent to some genes excluding exons and introns.

[00072] An object of the invention is to provide a method of identifying DNA sequences that control the expression of different collections of genes comprising, detecting, by computer, one or more pairs of non-adjacent DNA sequences to which are bound to two RNA sequences.

[00073] An object of the invention is to provide a method of identifying DNA sequences that control the expression of different collections of genes in a genome comprising detecting changes in connectron behavior in the genome.

[00074] An object of the invention is to provide a method of modifying the expression of different gene collections in a genome, comprising detecting changes in connectron behavior as a result of an exogenous stimulus.

[00075] An object of the invention is to provide a method of detecting where and when new genes are being integrated into a host genome comprising detecting the connectrons in said host genome.

[00076] An object of the invention is to provide a method of detecting the expression effect of different gene collections

in a given body comprising detecting the back and forth flow of connectrons between the chromosomes thereof.

[00077] An object of the invention is to provide a method of modifying a given body comprising modifying the connectron organization therein.

[00078] An object of the invention is to provide a method of detecting connectron control and target sequences in a given genome comprising:

- determining the base composition of said genome,
- determining one or more sites of control sequence organization, and/or
- determining one or more sites of target application.

[00079] An object of the invention is to provide a method of determining the response of a cell in any tissue to changes in the cell's environment and/or genetic composition comprising providing a complete genomic DNA sequence for the organism and determining the effect of changes in connectrons due to application of a given exogenous stimulus to the genome.

[00080] An object of the invention is to provide a method of determining in prokaryotes, archea, single-celled eukaryotes and multi-celled eukaryotes, the tetradic relationship $T1=C1$ and $T2=C2$ where $T1$ and $T2$ are DNA sequences 20 or more bases in length, where the $C1$ sequence is adjacent to the $C2$ sequence, where the $T1$ and $T2$ sequences are on the same chromosome, and where the $C1/C2$ sequences are on the same chromosome as $T1$ and $T2$ or where the $C1/C2$ sequences are on a chromosome different from $T1$ and $T2$, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 40 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart.

[00081] An object of the invention is to provide a method of determining in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes, the connectron relationship that permits many different C1/C2 short loops to control the existence of a T1-T2 long loop and wherein said C1/C2 short loops can be on the same chromosome or on different chromosomes from the T1-T2 long loop, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 40 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart.

[00082] An object of the invention is to provide a method of determining in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes, the connectron relationship that permits one C1/C2 short loop to control the existence of many T1-T2 long loops, the C1/C2 short loop can be on the same chromosome or on different chromosomes from the T1-T2 long loops, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 40 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart.

[00083] An object of the invention is to provide a method of determining in the connectron relationships between prokaryotes and their plasmids wherein said connectrons implement a control mechanism between the two genomes that makes it possible from them to form a symbiotic relationship, and in the case of *D. radiodurans* the relationship is not symmetric, and the *D. radiodurans* genome sends C1/C2 short loops to the MP1 plasmid, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 40 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart.

[00084] An object of the invention is to provide a method of determining that connectron relationships that exist in plant and higher animals.

[00085] An object of the invention is to provide a method of determining in prokaryotes, archea, single-celled eukaryotes and multi-celled eukaryotes, the connectron relationship that permits one C1/C2 short loop to control the existence of one or more T1-T2 long loops without being subject to any expression controls other than those of the gene to which the C1/C2 is 3'UTR, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 540 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart,

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart, and

3'UTR - untranslated 3' end of an mRNA is beyond the end of the last exon, a stop codon in the mRNA causes the ribosome to stop the translation of mRNA into protein.

[00086] An object of the invention is to provide a method of determining in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes, the connectron relationship that permits one C1/C2 short loop to control the existence of one or more T1-T2 long loops such that this C1/C2 short loop is itself subject to expression control by another T1-T2 long loop which surrounds it, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 540 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart.

[00087] An object of the invention is to provide a method of determining in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes, the connectron relationship that permits one C1/C2 short loop to control the existence of the T1-T2 long loop that surrounds it, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 40 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same

chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart.

[00088] An object of the invention is to provide a method of determining the connectron relationships that do not have any genes within the T1-T2 long loop, wherein:

T1 sequence is any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, and the T2 or T1 sequences must be between about 1kb and 105kb apart.

[00089] An object of the invention is to provide a method of determining the geneless connectron relationship where one C1/C2 short loop controls the existence of many geneless T1-T2 long loops, wherein:

C1 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C2 sequence must occur in the same chromosome as the C1 sequence,

C2 sequence - any positive or negative strand DNA sequence of 20 bases or more, the C1 sequence must occur in the same chromosome as the C2 sequence,

C1/C2 - any positive or negative strand DNA sequence of 40 or more bases such that the C1 sequence is adjacent to the C2 sequence,

T1 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T2 sequence, the T1 and T2 sequences must be between about 1kb and 105kb apart, and

T2 sequence - any positive or negative strand DNA sequence of 20 bases or more that is on the same chromosome as the T1 sequence, the T2 or T1 sequences must be between about 1kb and 105kb apart.

Description of the Drawings and Tables

[00090] The above and other objects, advantages and features of the invention will become more apparent when considered with the following specification and accompanying drawings and tables wherein:

[00091] Figure 1 DNA is structured in six levels of increasing condensation. Double stranded DNA is level 1. Two turns of DNA are wrapped about each chromatin particle at level 2. The chromatin particles which each containing 200 base pairs form into 30 nm particles at level 3. The 30 nm particles form into large loops with an approximate dimension of 300 nm at level 4. Metaphase chromosomes form a condensed structure with an approximate dimension of 700 nm at level 5. An entire metaphase chromosome has a width of approximately 1400 nm at level 6. The large loops at level 4 of the DNA structuring are thought to have between 20,000 (20 kb) and 100,000 (100 kb) base pairs.

The Molecular Biology of the Cell by Alberts, Bray, Lewis, Raff, Roberts and Watson, 3rd. ed. , Garland Publishing, Inc., New York, 1994, p. 354

[00092] Figure 2 (a) Chromatin DNA forms into a six-fold symmetry 30nm particles.

(b) The six-fold symmetry 30nm particles form a linear chain with a varying number of repeat units.

The Molecular Biology of the Cell by Alberts, Bray, Lewis, Raff, Roberts and Watson , 3rd.

ed. , Garland Publishing, Inc., New York, 1994,
p. 345

[00093] Figure 3 Long loops of 30nm particles are thought
to be closed at the bottom of the loop by proteins.

The Molecular Biology of the Cell by Alberts,
Bray, Lewis, Raff, Roberts and Watson, 3rd. ed.,
Garland Publishing, Inc., New York, 1994, p. 348

[00094] Figure 4(a) Transcription and Editing. (b) Movement of
the RNA through the Nucleus. (c) Connectron Formation

[00095] Figure 5 Overview of schematic organization of a
typical transcriptionally active chromosomal loop.

[00096] Table 1 Connectron Properties for Prokaryotic, Archea
and Eukaryotic Genomes

[00097] Table 2 Yeast Inter-Chromosomal Connectron Distribution

[00098] Figure 6 Genome size plotted as a log-log function
of the Number of Connectrons

[00099] Figure 7 Number of Sequence Instances plotted as a function of the Number of
Fragments

[000100] Figure 8 Level 0 - The overall view of the
algorithm

[000101] Figure 9 Level 1 - Process Flow of the Algorithm

[000102] Figure 10 Level 2a - two pages - Process Genome into Blocking Fragment File

[000103] Figure 11 Level 2b - three pages - Compute the Connectrons for a Genome

[000104] Figure 12 Level 2c - two pages - Analyze Possible Connectrons

[000105] Figure 13 Level 3a - Setup Genome Usage Memory

[000106] Figure 14 Level 3b - Find DBP-Size Blocking File for T1-Window

[000107] Figure 15 Level 1 - Find DBP-Size Blocking File for T2-Window

[000108] Figure 16 Level 2a - two pages - Find C1/C2 Entries

[000109] Figure 17 Level 2b - two pages - Scan Genome Usage Memory for Potential Connectrons

Description of the Invention

[000110] A connectron is a relationship among four DNA sequences. Each sequence must be at least 20 bases long. There is a report by Sharp and Zamore [3] that RNA sequences of "about length 25" are important as sources of RNAi. 27 bases were actually used as the minimum length of each of the sequences. The T1 sequence is on one strand of some chromosome in a genome. The T2 sequence is on the same strand of the same chromosome as the T1 sequence. The T1 and T2 sequences (which are each at least 20 bases in length) must be at least 5,000 bases distant from each other but they can not be more than 105,000 bases distant from each other. The C1 sequence and the C2 sequence (which are each at least 20 bases in length) are adjacent to each other on some strand of some chromosome in the genome. The C1/C2 sequences - called the "short loop" - can be on the same strand as the T1 and T2 sequences or they can be on the opposite strand. The C1/C2 sequences of the short loop can be on the same chromosome as the T1 and T2 sequences but they can also be on a different chromosome in the genome. When a genome has only one chromosome, then the point is moot. Many genomes, of course, have several chromosomes. The C1 sequence is identical to the T1 sequence and the C2 sequence is identical to the T2 sequence.

[000111] The C1/C2 sequence must be on the same strand as a gene, either be directly adjacent to the gene (i.e. a gap of 0 bases) for prokaryotic genomes or at this time be within 10,000 bases for eukaryotic genomes. The size of the gap between the end of the gene and the beginning of the C1/C2 sequence is a variable. The C1/C2 short loop is expressed as the 3'UTR (Un-Translated Region) of the gene. In the case of

prokaryotic genes that do not normally have introns, the whole mRNA becomes the active species for connectron formation. In the case of eukaryotic genes, the whole transcript is the active species for connectron formation upon editing of the transcript to eliminate the introns. The whole transcript then can move about in the cytoplasm of prokaryotic cells or the nucleus of eukaryotic cells. Since the C1 sequence is equivalent to the T1 sequence and the C2 sequence is equivalent to the T2 sequence, the C1 RNA can form a Hoogsteen triple-stranded RNA/DNA/DNA helix with the double-stranded T1 sequence. Similarly the C2 RNA can form a Hoogsteen triple-stranded RNA/DNA/DNA helix with the double-stranded T2 sequence. Because the C1 sequence and the C2 sequence are adjacent to each other, the C1/T2 RNA/DNA/DNA Hoogsteen triple helix is brought into physical adjacency to the C2/T2 RNA/DNA/DNA Hoogsteen triple helix. RNA/DNA/DNA hybrid helices are the most stable form of triple helix. RNA double helices, DNA double helices, RNA triple helices and DNA triple helices are all significantly less stable than a RNA/double-stranded DNA triple helix. The stable physical adjacency of the two triple-stranded Hoogsteen helices ensures that the long loop of double-stranded DNA between the T1 sequence and the T2 sequence can then be structured into 30 nm chromatin particles as shown in level 4 of figure 1. The genes on either strand of the DNA between the T1 sequence and the T2 sequence when they are structured into the 30 nm chromatin particles are not open to promotion and expression.

[000112] The tetradic relationship between the T1 and T2 sequences that form the long loop and the C1/C2 sequences that form the short loop are called a connectron. The name "connectron" was suggested by J. David Rawn Ph.D. of Towson University. A connectron is possible if the T1, T2, C1 and C2 sequences exist. A connectron is real if the C1/C2 short loop

sequence is adjacent to an expressible gene. If the expression of the adjacent gene is inside one or more T1 - T2 long loops then this connectron is said to be transient. If the adjacent gene is not inside any possible T1-T2 long loop then the connectron is said to be permanent. If a connectron is inside of a T1-T2 long loop that has the same sequences (i.e. T1 is really equal to C1 and T2 is really equal to C2) then the connectron is said to be self-limiting. This is true because once the C1/C2 sequence is expressed it forms the T1-T2 long loop that then shuts off the expression of the gene adjacent to the C1/C2 sequence. Self-limiting connectrons can also be called "spike" connectrons since they generate a short-duration spike of the C1/C2 short loop sequence. If a T1-T2 long loop does not contain any genes but it contains C1/C2 short loop sequences then this type of connectrons is said to be geneless. The C1/C2 short loops within a geneless T1-T2 long loop can, of course, control the expression of genes.

[000113] The physical existence and lifetimes of the connectrons must be proved by molecular biological experimentation. This physical experimental process, however, is logically quite separate from the computational experimentation that have been conducted from June of 1999 to May of 2001. The computational search for the existence of connectrons has been extremely positive. These computations have shown that connectrons exist in prokaryotes, in archea, between prokaryotes and their plasmids, in single-celled eukaryotes, in multi-celled eukaryotes, in plants, in higher animals and in humans. All of these features and properties are described in the claims section that follows.

[000114] The connectron invention is very powerful. It depends only on sequence equivalency. The minimum length of

the four sequences seems to be about 20 bases. In the calculations shown in this patent application, 27 bases have been used as a minimum. The Nature News Feature [1] says that other scientists have found RNA sequences of length about 25 that have interesting gene silencing properties. The Nature article does not give any mechanism. Because of my algorithm and its use on a variety of genomes, this patent application provides the computational proof that a particular mechanism is highly probable. The connectron invention provides an explanation for how communication occurs with a chromosome as well as between chromosomes in genomes that have more than one chromosome. Since each T1-T2 long loop can contain one or more genes, the connectron invention provides a mechanism for turning on and turning off sets of genes simultaneously. In time, the connectron invention will provide an explanation for how differentiation of how one cell's behavior differs from the behavior of another adjacent cell. It is already clear from the computational experiments that have been made on *S. cerevisiae*, *C. elegans* and *D. megalomaster* that the number of geneless connectrons increases dramatically as evolution proceeds from single-celled eukaryotes (i.e. *S. cerevisiae*) to 1,000 cell eukaryotes (i.e. *C. elegans*) to visible creatures (i.e. *D. megalomaster*). The extension of this evolutionary progress to plants (i.e. *A. thaliana*) for which only three chromosomes are sequenced and humans (i.e. *H. sapiens*) for which only one chromosome is completely sequenced. Although the complete human genome was published in Nature and Science in February of 2001, the NIH-sponsored genomic sequencing results are available for about 1/3 of the bases in the whole genome. The human genomic sequence determined by Celera Genomics, Inc. is available only by subscription. Table 1 shows how the genome size, the number of genes, the number of gene-containing and geneless connectrons and the percentage of genes controlled are related in many different genomes.

[000115] The C1/C2 short loops originate on one chromosome. The T1-T2 long loops can be on the same or different chromosomes. Table 2 which is for yeast (*S. cerevisiae*) is a square matrix of how many C1/C2 short loops on a given chromosome are sent to form T1-T2 long loops on other chromosomes. The diagonal of this matrix shows that many chromosomes send connectrons to themselves. The striking feature of this particular table is that chromosome 6 only sends connectrons to chromosome 12 but that it receives connectrons from chromosomes 4,5,7,10,12,13,15 and 16.

[000116] Any tetrad of connectron sequences (i.e. the T1, T2, C1 and C2 sequences) as well as the fact of the adjacency of the C1/C2 short loop sequence to the transcribing gene can be patented because the content of matter and the utility can be exactly described. The utility of a connectron is that the T1-T2 long loop shuts off the expression of the genes that lie between the T1 sequence and the T2 sequence. In the case of geneless connectrons, the utility is of a higher level in that the C1/C2 short loops contained in the higher-level geneless T1-T2 long loop, eventually form other lower-level T1-T2 long loops around a set of genes.

[000117] The invention of connectrons comes at a particularly important time in biological discovery. The geneless connectrons make a many-to-many hierarchical control mechanism possible. It is already clear from the determination of the connectrons for *C. elegans* and *D. megalomaster* that there are as many or more geneless connectrons than there are genes. It has been clear for some time that the number of genes in a genome is not particularly correlated with the size of the genome. Figure 6 shows that the size of a genome is roughly linearly correlated with the number of connectrons.

[000118] The connectron invention can be used to generate a model of behavior in any cell. The simulation of connectron behavior in different genomes will be the subject of another patent application.

[000119] The connectron invention provides for a rational exploitation of the information contained in the raw genomic DNA sequence by forming a hierarchy of relationships between geneless connectrons, transient connectrons, permanent connectrons, self-limiting connectrons and the expression of genes.

Detailed Description of the Invention

[000120] The algorithm for the determination of connectrons in any genome or any genome fragment is represented in the following flow diagrams. The Level 0 diagram in figure 8 shows the general relationships in a digital computer. The central processor of the digital computer uses the computer program to take genome descriptors, the genomic DNA sequences and the tables of gene features to produce a file of blocking fragments and a file of the optimal connectrons for the genome. The printer serves to make hard copies of the files and this patent application. The level 1 diagram in figure 9 shows the three essential steps in the determination of connectrons. The genome is first processed into a blocking fragment file. Then the blocking fragments are used to compute the connectrons for the genome. Finally the potential connectrons are analyzed to determine if the C1/C2 sequences are in the 3'UTR of a gene. The level 2a diagram in figure 10 shows the steps required for the processing of the genome into a file of blocking fragments. The genomic DNA sequence is decomposed into 27-base frames for both the positive and negative strands. These fragments are written to the unsorted fragment file. The fragment file is then sorted is then read and formed into groups of equivalent sequences. The (.blk) file contains the sequence and a pointer to the (.gpitr) file which contains the pointers to the position of the fragments in the genomes. The position in the genome includes the chromosome number, the position in the chromosome and the strand (i.e. positive and negative). A sample of these files follows

Sample of the (.blk) file for *S. cerevisiae*

27-base fragment (.gpstr) file	Number of instances	Pointer to
11111111111111111111111111111111	0	1
111111123244233313332443414	1	2
111111141113443133314333341	2	4
111111232442333133324434141	1	5
111111323311133323144423444	2	7
111111332213331341414443413	2	9
111111333444112343412323243	1	10
111111333444113343412323243	9	19
111111411134431333143333414	2	21
111111443223134142124434124	2	23
111112223234344444443144442	2	25
111112244123441122214421213	8	33
111112311241114344334134431	2	35
111112324423331333244341414	1	36
111112344232231344242234342	1	37
111112433444244421144134211	1	38
111112444311313442332142224	1	39
111113131241131114424413231	1	40
111113143332344311113133411	1	41
111113233111333231444234441	2	43

In fragments above 1=G, 2=C, 3=A, 4=T

Sample of the (.gpstr) file for *S. cerevisiae*

There are 16 chromosomes in *S. cerevisiae*

Item	Chromosome	Position in Chromosome	Direction
1	0	0	0
2	4	11137	1
3	12	467619	1
4	12	458482	1
5	4	11138	1
6	12	465759	2
7	12	456622	1
8	1	219366	1
9	8	539978	1
10	14	522451	1
11	4	1099073	1
12	4	1210003	1
13	7	539068	1

14	12	654136	1
15	12	596455	1
16	15	121016	1
17	15	598127	2
18	16	847724	1
19	16	59765	1
20	12	467620	1
21	12	458483	1
22	12	461657	1
23	12	452520	1
24	13	838006	1
25	15	288270	1
26	4	83593	1
27	4	992867	1
28	6	162265	1
29	7	845687	1
30	10	531560	2
31	15	282208	1
32	16	860418	1
33	16	572308	1
34	12	465992	1
35	12	456855	1
36	4	11139	1
37	8	89343	1
38	4	10302	1
39	1	19894	2
40	16	9311	1
41	10	735203	1
42	12	465760	1
43	12	456623	1

In direction column above 1=positive strand, 2=negative strand

[000121] The level 2b diagram in figure 11 shows the computation of the connectrons. The genome descriptors consist of the number and length of the chromosomes. The algorithm uses an array that represents several facts about each base position in the genome. The level 3a diagram in figure 13 shows the setup of the Genome-Usage memory. The gene features are used to prevent the region of the genome that codes for proteins from being used for the connectron sequences (i.e. the T1s, the T2s, the C1s and the C2s). In the level 2a diagram of figure 10, the algorithm steps through each chromosome and within each chromosome through each base

position looking for acceptable T1-windows of 27 bases. A T1-window can be used to form a connectron relationship if there are two or more instances of this fragment in the blocking fragment file. The computation in the level 3b diagram of figure 14 determines if the T1-window is acceptable or not. Once an acceptable T1-window is found, the algorithm (in the level 2a diagram of figure 10) looks for acceptable T2-window positions that lie between 5,000 and 105,000 bases from the T1-window. The computation for determining acceptable T2-window positions is done in the level 3c diagram of figure 15. Once a pair of T1 and T2 window positions are found, the algorithm looks among the instances of these T1 and T2 sequences for a pair of sequences C1 and C2 that lie within 200 bases of each other on the same chromosome. The computation for determining acceptable C1/C2 windows is shown in the level 3d diagram in figure 16. In the level 3e diagram of figure 17 the Genome-Usage memory is scanned for the Possible-Connectrons. In the level 2c diagram of figure 12 the Possible-Connectrons are scanned to determine if the C1/C2 sequences are within the Gap-Distance of a gene on either the positive or the negative strands. The Real-Connectrons are then written out in several different files including the descriptions in the claims section.

Examples

[000122] The algorithm for the determination of optimal connectrons has been applied to a number of different publicly available genomes. The connectron is a tetradic relationship between four sequence elements - T1, T2, C1 and C2. The claims presented in this section are written by the program NearGene that implements the flow diagram Level 2c of figure 12. The examples are written a uniform type of English. Each example contains some or all of the following elements

```
-----
Name of genome
Description of T1
Length of T1-T2 loop
The chromosome on which the T1-T2 loop exists
The identifier number within the genome of the T1
sequence
The T1 sequence
Description of T2
The identifier number within the genome of the T2
sequence
The T2 sequence
A list of genes whose expression is controlled by the
T1-T2 loop
The common names of the genes as obtained from the NCBI
gene feature file (.ptt)
A list of C1/C2 short loops whose expression if
controlled by the T1-T2 loop
The chromosome on which the C1/C2 short loop exists
The common name of the gene which expresses the C1/C2
short loop as an RNA
The sequence of the C1/C2 short loop
```

A list of C1/C2 short loops that control the formation
of the T1-T2 loop

The chromosome on which the C1/C2 short loop exists

The common name of the gene which expresses the C1/C2
short loop as an RNA

The sequence of the C1/C2 short loop

The match between the C1/C2 sequence and the T1
sequence

The match between the C1/C2 sequence and the T2
sequence

[000123] The uniform descriptions make it possible to rapidly
comprehend the specifics in each example.

[000124] When a sequence element is very long a series of
four dots has been inserted between the beginning and ending
sequence groups. A variable number of bases have been
deleted.

1. Connectrons occur in prokaryotes, archea, single-celled eukaryotes and multi-celled eukaryotes.
2. Many Connectrons control the expression of one set of genes in prokaryotes, archea, single-celled eukaryotes and multi-celled eukaryotes.
3. One connectron controls the expression of many sets of genes in prokaryotes, archea, single-celled eukaryotes and multi-celled eukaryotes.
4. Connectrons occur between prokaryotes and their plasmids.
5. Connectrons occur in plants and higher animals
6. Permanent connectrons exist in prokaryotes, archea,
- single-celled eukaryotes and multi-celled eukaryotes.
7. Transient connectrons exist in prokaryotes, archea, single-celled eukaryotes and multi-celled eukaryotes.
8. Self-limiting connectrons occur in prokaryotes, archea, single-celled eukaryotes and multi-celled eukaryotes
9. Geneless connectrons exist in single-celled and multi-celled eukaryotes

10. One connectron controls many geneless
connectrons in single-celled and multi-celled
eukaryotes

1. Connectrons occur in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[000126] Connectrons exist as tetradic relationships where the sequence T1 is equivalent to the sequence C1 (written T1=C1) and where the sequence T2 equals the sequence C2 (written T2=C2) where T1 and T2 are DNA sequences 20 or more bases in length, where the C1 sequence is adjacent to the C2 sequence, where the T1 and T2 sequences are on the same chromosome, and where the C1/C2 sequences are on the same chromosome as T1 and T2 or where the C1/C2 sequences are on a chromosome different from T1 and T2. The connectron relationship has been found to exist in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[000127] Example of a prokaryote connectron - E. coli

[000128] In this example the existence of the T1-T2 (3197-3308) long loop is controlled by three C1/C2 short loops (3307, 3432 and 2218). The T1-T2 long loop controls the expression of 64 genes on chromosome 1 in addition to six C1/C2 (3204, 3206, 3223, 3228, 3301 and 3327) short loops. The C1/C2 short loop 3327 lies outside the range of the T1-T2 long loop (3197-3308) but this C1/C2 is expressed as a 3'UTR to the gene hemG that is within the range of the T1-T2 long loop.

```

3307 Chromosome 1
3432 Chromosome 1
2218 Chromosome 1
|
*-----*
|           Chromosome 1           |
3197                                     3308
|           3204   3206               |
|           3224   3228               |
|           3301   3327               |

```

[000129] Connectron control elements for chromosome 1 of the E. coli genome

[000130] A double stranded DNA loop of length 93.542 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 3197. This T1 control element has the DNA sequence

[000131] Seq. Id. = 1 Position = 1 to 175

[000132] AAAAAATGCGCGGTCAGAAAATTATTTTAAATTTCTTGTCTGTCAGGCCGGAA
TAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCG
GGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAA

[000133] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3308. This T2 control element has the DNA sequence

[000134] Seq. Id. = 2 Position = 1 to 175

[000135] TAAATTTCTTGTCTGTCAGGCCGGAATAAATCCCTATAATGCGCCACCACTGA
CACGGAACAACGGCAAACACGCCGCCGGGTCAGCGGGGTTCTCCTGAGAACTCCGGCAGAGA
AAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTATTATGCACACCCCGCGCCGCT

[000136] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

rrsC	gltU	rrlC	rrfC	aspT
trpT	yifA	yifE	yifB	ilvL
ilvG_1	ilvM	ilvE	ilvD	ilvA
ilvY	ilvC	ppiC	b3776	rep
gppA	rhlB	trxA	rhoL	rho
rfe	wzzE	wecB	rffH	wecD
wecE	wzxE	yifM_2	wecG	yifK
argX	hisR	leuT	proM	aslB
aslA	hemY	hemX	hemD	cyaA
cyaY	b3808	dapF	uvrD	b3814
corA	yigF	yigG	rarD	yigI
pldA	recQ	yigJ	yigK	pldB
yigL	yigM	metR	metE	ysgA
udp	yigN	ubiE	yigP	b3836
yigU	yigW_1	rfaH	yigC	ubiB
fadA	fadB	pepQ	trkH	hemG

[000137] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000138] A C1/C2 short loop on chromosome 1 whose identifier is 3204 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrsC and has the DNA sequence

[000139] Seq. Id. = 3 Position = 1 to 186

[000140] GATGTGCCCAGATGGGATTAGCTAGTAGGTGGGGTAACGGCTCACCTAGGCG
ACGATCCCTAGCTGGTCTGAGAGGATGACCAGCCACACTGGAAGTCTGAGACACGGTCCAGACT

CCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCG
CGTGTATGAA

[000141] A C1/C2 short loop on chromosome 1 whose identifier is 3206 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrsC and has the DNA sequence

[000142] Seq. Id. = 4 Position = 1 to 186

[000143] GTCCCCTTCGTCTAGAGGCCAGGACACCGCCCTTTCACGGCGGTAACAGGG
GTTCGAATCCCCTAGGGGACGCCACTTGCTGGTTTGTGAGTGAAAGTCACCTGCCTTAATAT
CTCAAACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAAGCTG
AAAATTGAAA

[000144] A C1/C2 short loop on chromosome 1 whose identifier is 3223 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrlC and has the DNA sequence

[000145] Seq. Id. = 5 Position = 1 to 186

[000146] GCTGAAGTAGGTCCCAAGGGTATGGCTGTTTCGCCATTTAAAGTGGTACGCGA
GCTGGGTTTAGAACGTCGTGAGACAGTTCGGTCCCTATCTGCCGTGGGCGCTGGAGAACTGA
GGGGGGCTGCTCCTAGTACGAGAGGACCGGAGTGGACGCATCACTGGTGTTTCGGGTGTGCAT
GCCAATGGCA

[000147] A C1/C2 short loop on chromosome 1 whose identifier is 3225 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrlC and has the DNA sequence

[000148] Seq. Id. = 6 Position = 1 to 144

[000149] AAACAGAATTTGCCTGGCGGCCGTAGCGCGGTGGTCCCACCTGACCCCATGC
CGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTA
GGGAACTGCCAGGCATCAAATTAAGCAGTA

[000150] A C1/C2 short loop on chromosome 1 whose identifier is 3228 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrfC and has the DNA sequence

[000151] Seq. Id. = 7 Position = 1 to 112

[000152] GGTCATAAAACCGGTGGTTGTAAAAGAATTCGGTGGAGCGGTAGTTCAGTCG
GTTAGAATACCTGCCTGTCACGCAGGGGGTCGCGGGTTCGAGTCCCGTCCGTTCCGCCAC

[000153] A C1/C2 short loop on chromosome 1 whose identifier is 3301 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ubiB and has the DNA sequence

[000154] Seq. Id. = 8 Position = 1 to 57

[000155] TTATCGTGCCTACAAATAGTCCGAACCGTAGGCCGGATAAGGCGTTTACGCC
GCATC

[000156] A C1/C2 short loop on chromosome 1 whose identifier is 3307 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene fadA and has the DNA sequence

[000157] Seq. Id. = 9 Position = 1 to 56

[000158] TGCCGGATGCGGCGTAAACGCCTTATCCGGCCTACGGTTCGGACTATTTGTA
GGCA

[000159] A C1/C2 short loop on chromosome 1 whose identifier is 3327 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene hemG and has the DNA sequence

[000160] Seq. Id. = 10 Position = 1 to 347

[000161] AAAAAATGCGCGGTCAGAAAATTATTTTAAATTCCTCTTGTCAGGCCGGAA
TAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCG
GGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAG
GCGTATTATG...CCCGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCT
TCGGGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAACCGT
AGGGGAACCTGCGGTTGGATCACCTCCTTACCTTAAAGAAGCGTTCTTTG

[000162] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000163] A C1/C2 short loop on chromosome 1 whose identifier is 3307 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene hemG and has the DNA sequence

[000164] Seq. Id. =11 Position = 1 to 347

[000165] AAAAAATGCGCGGTCAGAAAATTATTTTAAATTCCTCTTGTCAGGCCGGAA
TAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCG
GGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAG

GCGTATTATG...CCCGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCT
TCGGGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAACCGT
AGGGGAACCTGCGGTTGGATCACCTCCTTACCTTAAAGAAGCGTTCTTTG

[000166] The match between the T1 sequence and the C1/C2
sequence is

[000167] Seq. Id. = 11 Position = 1 to 175

[000168] AAAAAATGCGCGGTCAGAAAATTATTTTAAATTCCTCTTGTGTCAGGCCGGAA
TAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCCGCGGGTCAGCG
GGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAA

[000169] The match between the T2 sequence and the C1/C2
sequence is

[000170] Seq. Id. = 11 Position = 28 to 202

[000171] TAAATTCCTCTTGTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGA
CACGGAACAACGGCAAACACGCCCGCGGGTCAGCGGGGTTCTCCTGAGAACTCCGGCAGAGA
AAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTATTATGCACACCCCGCGCCGCT

[000172] A C1/C2 short loop on chromosome 1 whose identifier
is 3432 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene btuB and has the DNA sequence

[000173] Seq. Id. = 12 Position = 1 to 335

[000174] TGCGCGGTCAGAAAATTATTTTAAATTCCTCTTGTGTCAGGCCGGAATAACTC
CCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCCGCGGGTCAGCGGGGTTCT
CCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTAT
TATGCACACC...ACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCGGGA

GGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAACCGTAGGGGA
ACCTGCGGGTTGGATCACCTCCTTACCTTAAAGAAGCGT

[000175] The match between the T1 sequence and the C1/C2 sequence is

[000176] Seq. Id. = 12 Position = 1 to 169

[000177] TGC GCGGTCAGAAAATTATTTTAAATTCCTCTTGT CAGGCCGGAATAACTC
CCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCGGGGTTC
TCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAA

[000178] The match between the T2 sequence and the C1/C2 sequence is

[000179] Seq. Id. = 12 Position = 22 to 196

[000180] TAAATTCCTCTTGT CAGGCCGGAATAACTCCCTATAATGCGCCACCACTGA
CACGGAACAACGGCAAACACGCCGCCGGGTCAGCGGGGTTCCTGAGAACTCCGGCAGAGA
AAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTATTATGCACACCCCGCGCCGCT

[000181] A C1/C2 short loop on chromosome 1 whose identifier is 2218 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene clpB and has the DNA sequence

[000182] Seq. Id. = 13 Position = 1 to 72

[000183] CTTGT CAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
CGGCAAACACGCCGCCGGGC

[000184] The match between the T1 sequence and the C1/C2 sequence is

[000185] Seq. Id. = 13 Position = 1 to 72

[000186] CTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
CGGCAAACACGCCGCCGGGC

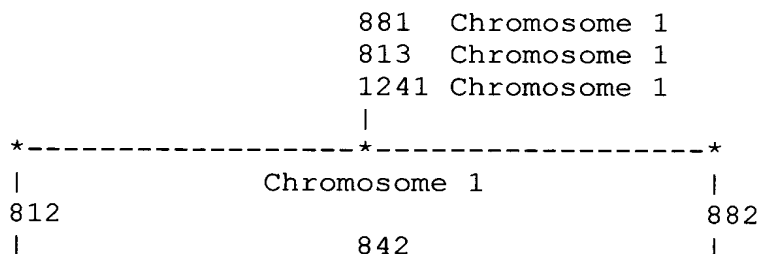
[000187] The match between the T2 sequence and the C1/C2
sequence is

[000188] Seq. Id. = 13 Position = 1 to 71

[000189] CTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
CGGCAAACACGCCGCCGGG

[000190] Example of an archea connectron - *H. pylori*

[000191] In this example the existence of the T1-T2 (812-
882) long loop is controlled by three C1/C2 short loops (881,
813 and 1214). The T1-T2 long loop controls the expression of
54 genes on chromosome 1 in addition to one C1/C2 (843) short
loop.



[000192] Connectron control elements for chromosome 1 of *H.*
pylori genome

[000193] A double stranded DNA loop of length 96.385 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 812. This T1 control element has the DNA sequence

[000194] Seq. Id. = 14 Position = 1 to 43

[000195] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCA

[000196] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 882. This T2 control element has the DNA sequence

[000197] Seq. Id. = 15 Position = 1 to 43

[000198] TAGCGGAACTAAAGCATTTCATCCCAAACACTAAAGATATTTGG

[000199] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

HP0999	HP1000	HP1001	HP1002	HP1003
HP1005	HP1006	HP1008	HP1009	HPtRNA-Pro
HP1010	HP1011	HP1013	HP1015	HP1017
HP1018	HP1020	HP1021	HP1022	HP1023
HP1024	HP1025	HP1027	HP1028	HP1030
HP1031	HP1033	HP1034	HP1038	HP1039
HP1040	HP1041	HP1042	HP1043	HP1044
HP1045	HP1046	HP1051	HP1052	HP1055
HP1056	HP1058	HP1060	HP1065	HPtRNA-Ser
HP1066	HP1067	HP1069	HP1070	HP1074
HP1075	HP1076	HP1077	HP1078	HP1079
HP1080	HP1081	HP1083	HP1084	HP1085
HP1088	HP1091	HP1092	HP1093	HP1094
HP1095	HP1096			

[000200] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000201] A C1/C2 short loop on chromosome 1 whose identifier is 813 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene HP0998 and has the DNA sequence

[000202] Seq. Id. = 16 Position = 1 to 70

[000203] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAAGTAAAGCATTTCATCCCA
AACACTAAAGATATTTGG

[000204] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000205] A C1/C2 short loop on chromosome 1 whose identifier is 881 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene HP1096 and has the DNA sequence

[000206] Seq. Id. = 17 Position = 1 to 70

[000207] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAAGTAAAGCATTTCATCCCA
AACACTAAAGATATTTGG

[000208] The match between the T1 sequence and the C1/C2 sequence is

[000209] Seq. Id. = 17 Position = 1 to 36

[000210] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCA

[000211] The match between the T2 sequence and the C1/C2 sequence is

[000212] Seq. Id. = 17 Position = 28 to 70

[000213] TAGCGGAACTAAAGCATTCATCCCAAACACTAAAGATATTTGG

[000214] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000215] A C1/C2 short loop on chromosome 1 whose identifier is 813 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene HP0998 and has the DNA sequence

[000216] Seq. Id. = 18 Position = 1 to 70

[000217] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCATTCATCCCA
AACACTAAAGATATTTGG

[000218] A C1/C2 short loop on chromosome 1 whose identifier is 881 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene HP1096 and has the DNA sequence

[000219] Seq. Id. = 19 Position = 1 to 70

[000220] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCATTCATCCCA
AACACTAAAGATATTTGG

[000221] The match between the T1 sequence and the C1/C2 sequence is

[000222] Seq. Id. = 19 Position = 1 to 43

[000223] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCA

[000224] The match between the T2 sequence and the C1/C2 sequence is

[000225] Seq. Id. = 19 Position = 28 to 70

[000226] TAGCGGAACTAAAGCATTTCATCCCAAACACTAAAGATATTTGG

[000227] A C1/C2 short loop on chromosome 1 whose identifier is 1241 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene HP1535 and has the DNA sequence

[000228] Seq. Id. = 20 Position = 1 to 56

[000229] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCATTTCATCCCA
AACA

[000230] The match between the T1 sequence and the C1/C2 sequence is

[000231] Seq. Id. = 20 Position = 1 to 43

[000232] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCA

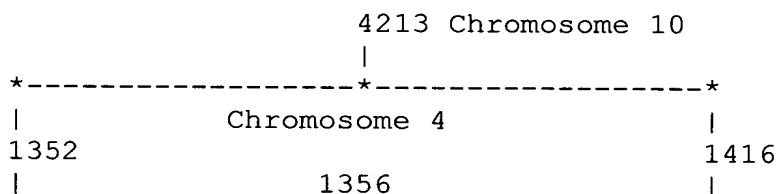
[000233] The match between the T2 sequence and the C1/C2 sequence is

[000234] Seq. Id. = 20 Position = 28 to 56

[000235] TAGCGGAACTAAAGCATTTCATCCCAAACA

[000236] Example of single-celled connectron - *S. cerevisiae*

[000237] In this example the existence of the T1-T2 (1352-1416) long loop on chromosome 4 is controlled by one C1/C2 short loop (4213) on chromosome 10. The T1-T2 long loop controls the expression of 34 genes on chromosome 4 in addition to one C1/C2 (1356) short loop.



[000238] Connectron control elements for chromosome 1 of *S. cerevisiae* genome

[000239] A double stranded DNA loop of length 68.908 kilobases on chromosome 4 is bounded on the left by a T1 sequence whose identifier is 1352. This T1 control element has the DNA sequence

[000240] Seq. Id. = 21 Position = 1 to 37

[000241] TTATGAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAA

[000242] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1416. This T2 control element has the DNA sequence

[000243] Seq. Id. = 22 Position = 1 to 362

[000244] ATTAGATCTATTACATTATGGGTGGTATGTTGGAATAAAAATCAACTATCATCTACTAACTAGTATTTACGTTACTAGTATATTATCATATACGGTGTAGAAAGATGACGCAAA TGATGAGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAACATATAAAACGATGATAATAATATTTATAGAATTGTGTAGAATTGCA GATTCCCTTTTATGGATTCCCTAAATCCTTGAGGAGAACTTCTAGTATATCTACATACCTAAT ATTATAGCCTTAATCACAATGGAATCCCAACAATTACATCAAAATCCACATTCTCTACAGTA

[000245] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

YDR170W-A	YDR171W	YDR172W	YDR173C	YDR174W
YDR175C	YDR176W	YDR177W	YDR178W	YDR179C
YDR179W-A	YDR180W	YDR181C	YDR182W	YDR183W
YDR184C	YDR185C	YDR186C	YDR187C	YDR188W
YDR189W	YDR190C	YDR191W	YDR192C	YDR193W
YDR194C	YDR195W	YDR196C	YDR197W	YDR198C
YDR199W	YDR200C	YDR201W	YDR202C	YDR203W
YDR204W	YDR205W	YDR206W	YDR207C	YDR208W
YDR209C	YDR210W			

[000246] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000247] A C1/C2 short loop on chromosome 4 whose identifier is 1356 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YDR170W-A and has the DNA sequence

[000248] Seq. Id. = 23 Position = 1 to 311

[000249] AATCACACTAATCATTCTGATGATGAACTCCCTGGACACCTCCTTCTCGATT
CAGGAGCATCACGAACCCTTATAAGATCTGCTCATCACATACACTCAGCATCATCTAATCCT
GACATAAACGTAGTTGATGCTCAAAAAAGAAATATACCAATTAACGCTATTGGTGACCTACA
ATTTCACTTCCAGGACAACACCAAAACATCAATAAAGGTATTGCACACTCCTAACATAGCCT
ATGACTTACTCAGTTTGAATGAATTGGCTGCAGTAGATATCACAGCATGCTTTACCAAAAAC
GTCTTAGAACG

[000250] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000251] A C1/C2 short loop on chromosome 10 whose identifier is 4213 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YJR029W and has the DNA sequence

[000252] Seq. Id. = 24 Position = 1 to 346

[000253] ATCTATTACATTATGGGTGGTATGTTGGAATAAAAATCCACTATCGTCTATC
AACTAATAGTTATATTATCAATATATTATCATATACGGTGTTAAGATGATGACATAAGTTAT
GAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCA
ATGAATATAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAG
ATTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAAT
ATTATAGCCTTTATCAACAATGGAATCCCAACAATTATCTCAACAT

[000254] The match between the T1 sequence and the C1/C2 sequence is

[000255] Seq. Id. = 24 Position = 111 to 147

[000256] TTATGAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAA

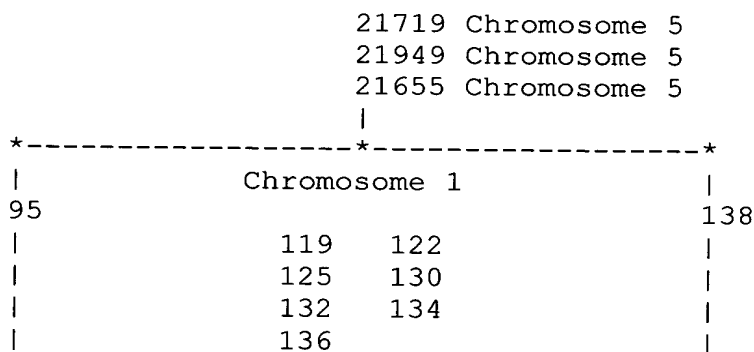
[000257] The match between the T2 sequence and the C1/C2 sequence is

[000258] Seq. Id. = 24 Position = 1 to 38

[000259] ATCTATTACATTATGGGTGGTATGTTGGAATAAAAATC

[000260] Example of a multi-celled connectron - C. elegans

[000261] In this example the existence of the T1-T2 (9-138) long loop on chromosome 1 is controlled by three C1/C2 short loops on chromosome 5 (21719, 21949 and 21655). The T1-T2 long loop controls the expression of four genes on chromosome 1 in addition to seven C1/C2 (119, 122, 125, 130, 132, 134 and 136) short loops.



[000262] A double stranded DNA loop of length 41.978 kilobases on chromosome 1 is bounded on the left by a T1 sequence

whose identifier is 95. This T1 control element has the DNA sequence

[000263] Seq. Id. = 25 Position = 1 to 55

[000264] CAGCACGTTCTTAACCATGCAAAATCAGTTGAGAACTCTGCGTCTCTTCTCC
CGC

[000265] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 138. This T2 control element has the DNA sequence

[000266] Seq. Id. = 26 Position = 1 to 36

[000267] ACTCTGCGTCTCTTCTCCCGCATTTTTTGTAGATCA

[000268] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

Y73A3A.1 Y73A3A.1 ZC123.3 ZC123.2

[000269] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000270] A C1/C2 short loop on chromosome 1 whose identifier is 119 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.3 and has the DNA sequence

[000271] Seq. Id. = 27 Position = 1 to 69

[000272] TTGAGAACTCTGCGTCTCAACTCCCGCATTTTTTGTAGATCTACGTAGATCA
AACCGAAATGGGACACT

[000273] A C1/C2 short loop on chromosome 1 whose identifier is 122 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.3 and has the DNA sequence

[000274] Seq. Id. = 28 Position = 1 to 89

[000275] GCACGGGGTTCTGGCCTTCCTCATTGAATTTTTCGCGCTCCATTGACAATCG
CCTGCCGGACAACGCGTGGGAAAGTCGTGTACTCCAC

[000276] A C1/C2 short loop on chromosome 1 whose identifier is 125 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.3 and has the DNA sequence

[000277] Seq. Id. = 29 Position = 1 to 89

[000278] ACGCGCCGTAAATCTACCCAGATATGGCCGAGCCAAAATGGCCTAGTTCGG
CAAACCTCTTTCATTTCAATTTATGAGGGAAGCCAGAA

[000279] A C1/C2 short loop on chromosome 1 whose identifier is 130 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.2 and has the DNA sequence

[000280] Seq. Id. = 30 Position = 1 to 121

[000281] CTCCCGCATTTTTTGTAGATCTACGTAGATCAAACCGAAATGAGGCACTTTC
TGAATCCACGAGCTAGGCTTAAGCTTAGGCTTAAGCTTAGGCCTTTCTCAGGCTTAGGCTT
AGGCTTA

[000282] A C1/C2 short loop on chromosome 1 whose identifier is 132 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.2 and has the DNA sequence

[000283] Seq. Id. = 31 Position = 1 to 190

[000284] GCTTATGCTTGGGCTTAGGCTTAGGCGTAGGCTTAGGCTTAGGCTTAGGCTT
ATGCTTAGACTTAGTCTCACTATCAGTCTTAGGCTTAGGCTTAGACTTAGGCTTAAGCTTAG
GCTTAAGCTTAGACTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGTTTGGGCTTAGGC
TTAGGCTTAACCTC

[000285] A C1/C2 short loop on chromosome 1 whose identifier is 134 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.2 and has the DNA sequence

[000286] Seq. Id. = 32 Position = 1 to 133

[000287] TCTGCGTCTTTTCTCCCGCATTTTTTGTAGATCTACGTAGATCAAACCGAAA
TGAGGCACTTTCTGAATCCACGAGCTAGGCTTAAGCTTAGGCTTAAGCTTAGGCCTTTTCTC
AGGCTTAGGCTTAGGCTTA

[000288] A C1/C2 short loop on chromosome 1 whose identifier is 136 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.2 and has the DNA sequence

[000289] Seq. Id. = 33 Position = 1 to 190

[000290] GCTTATGCTTGGGCTTAGGCTTAGGCGTAGGCTTAGGCTTAGGCTTAGGCTT
ATGCTTAGACTTAGTCTCACTATCAGTCTTAGGCTTAGGCTTAGACTTAGGCTTAAGCTTAG
GCTTAAGCTTAGACTTAGGCTTAGGCTTAGGCTTAGGCTTAGGCTTAGGTTTGGGCTTAGGC
TTAGGCTTAACCTC

[000291] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000292] A C1/C2 short loop on chromosome 5 whose identifier is 21719 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene C39F7.5 and has the DNA sequence

[000293] Seq. Id. = 34 Position = 1 to 65

[000294] ACGTTCTTAACCATGCAAAATCAGTTGAGAACTCTGCGTCTCTTCTCCCGCA
TTTTTTGTAGATC

[000295] The match between the T1 sequence and the C1/C2 sequence is

[000296] Seq. Id. = 34 Position = 1 to 51

[000297] ACGTTCTTAACCATGCAAAATCAGTTGAGAACTCTGCGTCTCTTCTCCCGC

[000298] The match between the T2 sequence and the C1/C2 sequence is

[000299] Seq. Id. = 34 Position = 31 to 65

[000300] ACTCTGCGTCTCTTCTCCCGCATTTTTTGTAGATC

[000301] A C1/C2 short loop on chromosome 5 whose identifier is 21949 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene F16B4.4 and has the DNA sequence

[000302] Seq. Id. = 35 Position = 1 to 95

[000303] ACCATGCAAAATCAGTTGAGAACTCTGCGTCTCTTCTCCCGCATTTTTTGTA
GATCTACGTAGATCAAGCCGAAATGAGACACTCTGACACCACG

[000304] The match between the T1 sequence and the C1/C2 sequence is

[000305] Seq. Id. = 35 Position = 1 to 42

[000306] ACCATGCAAAATCAGTTGAGAACTCTGCGTCTCTTCTCCCGC

[000307] The match between the T2 sequence and the C1/C2 sequence is

[000308] Seq. Id. = 35 Position = 22 to 63

[000309] ACTCTGCGTCTCTTCTCCCGCATTTTTTGTAGATC

[000310] A C1/C2 short loop on chromosome 5 whose identifier is 21655 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene C39F7.3 and has the DNA sequence

[000311] Seq. Id. = 36 Position = 1 to 61

[000312] AACCATGCAAAATCAGTTGAGAACTCTGCGTCTCTTCTCCCGCATTTTTTGT
AGATCTACG

[000313] The match between the T1 sequence and the C1/C2
sequence is

[000314] Seq. Id. = 36 Position = 1 to 36

[000315] AACCATGCAAAATCAGTTGAGAACTCTGCGTCTCTTCTCCCGC

[000316] The match between the T2 sequence and the C1/C2
sequence is

[000317] Seq. Id. = 36 Position = 23 to 57

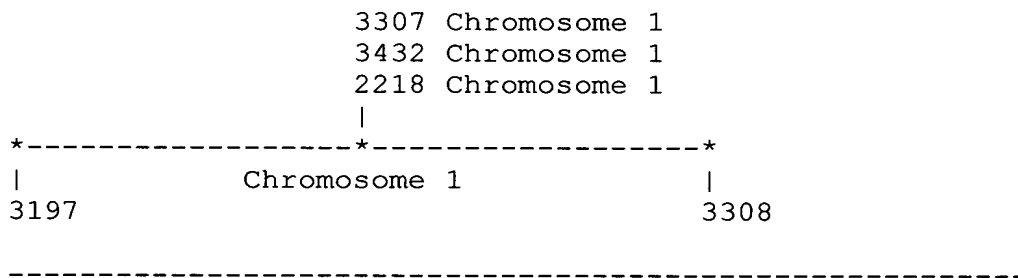
[000318] ACTCTGCGTCTCTTCTCCCGCATTTTTTGTAGATC

2. Many Connectrons control the expression of one set of genes in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[000319] Many different C1/C2 short loops can control the existence of one T1-T2 long loop. The C1/C2 short loops can be on the same chromosome or on different chromosomes from the T1-T2 long loop. This relationship is described as "many-to-one". This relationship exists in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes

[000320] Example of a many-to-one connectron in prokaryotes - E. coli

[000321] In this example the existence of the T1-T2 (3197-3308) long loop is controlled by three C1/C2 short loops (3307, 3432 and 2218).



[000322] A double stranded DNA loop of length 93.542 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 3197. This T1 control element has the DNA sequence

[000323] Seq. Id. = 37 Position = 1 to 175

[000324] AAAAAATGCGCGGTCAGAAAATTATTTTAAATTTCTCTTGTCAGGCCGGAA
TAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCG
GGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAA

[000325] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3308. This T2 control element has the DNA sequence

[000326] Seq. Id. = 38 Position = 1 to 175

[000327] TAAATTTCTCTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGA
CACGGAACAACGGCAAACACGCCGCCGGGTCAGCGGGGTTCTCCTGAGAACTCCGGCAGAGA
AAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTATTATGCACACCCCGCGCCGCT

[000328] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

rrsC	gltU	rrlC	rrfC	aspT
trpT	yifA	yifE	yifB	ilvL
ilvG_1	ilvM	ilvE	ilvD	ilvA
ilvY	ilvC	ppiC	b3776	rep
gppA	rhlB	trxA	rhoL	rho
rfe	wzzE	wecB	rffH	wecD
wecE	wzxE	yifM_2	wecG	yifK
argX	hisR	leuT	proM	aslB
aslA	hemY	hemX	hemD	cyaA
cyaY	b3808	dapF	uvrD	b3814
corA	yigF	yigG	rarD	yigI
pldA	recQ	yigJ	yigK	pldB
yigL	yigM	metR	metE	ysgA
udp	yigN	ubiE	yigP	b3836
yigU	yigW_1	rfaH	yigC	ubiB
fadA	fadB	pepQ	trkH	hemG

[000329] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000330] A C1/C2 short loop on chromosome 1 whose identifier is 3307 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene hemG and has the DNA sequence

[000331] Seq. Id. = 39 Position = 1 to 440

[000332] AAAAAATGCGCGGTCAGAAAATTATTTTAAATTTCTCTTGTCAGGCCGGAA
TAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCG
GGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAG
GCGTATTATG...GGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGTAATCGTGGA
TCAGAAATGCCACGGTGAATACGTTCCCGGGCCTTGACACACCGCCCGTCACACCATGGGAG
TGGGTTGCAAAAGAAGTAGGTAGCTTAACCTTCGGGAGGGCGCTTACCACTTTGTGATTCAT
GACTGGGGTGAAGTCGTAACAAGGTAACCGTAGGGGAACCTGCGGTTGGATCACCTCCTTAC
CTTAAAGAAGCGTTCTTTG

[000333] The match between the T1 sequence and the C1/C2 sequence is

[000334] Seq. Id. = 39 Position = 1 to 175

[000335] AAAAAATGCGCGGTCAGAAAATTATTTTAAATTTCTCTTGTCAGGCCGGAA
TAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCAGCG
GGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAA

[000336] The match between the T2 sequence and the C1/C2 sequence is

[000337] Seq. Id. = 39 Position = 28 to 192

[000338] TAAATTTCTCTTGTCTAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCTAGCGGGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTATTATGCACACCCCGCGCCGCT

[000339] A C1/C2 short loop on chromosome 1 whose identifier is 3432 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene btuB and has the DNA sequence

[000340] Seq. Id. = 40 Position = 1 to 335

[000341] TCGCGGTCAGAAAATTATTTTAAATTTCTCTTGTCTAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCTAGCGGGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTATTATGCACACC...ACACCATGGGAGTGGGTGCAAAAGAAGTAGGTAGCTTAACCTTCGGGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAACCGTAGGGGAACCTGCGGTTGGATCACCTCCTTACCTTAAAGAAGCGT

[000342] The match between the T1 sequence and the C1/C2 sequence is

[000343] Seq. Id. = 40 Position = 1 to 169

[000344] TCGCGGTCAGAAAATTATTTTAAATTTCTCTTGTCTAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCTAGCGGGGTTCTCCTGAGAACTCCGGCAGAGAAAGCAAAAATAAATGCTTGACTCTGTAGCGGGAA

[000345] The match between the T2 sequence and the C1/C2 sequence is

[000346] Seq. Id. = 40 Position = 22 to 196

[000347] TAAATTTCTCTTGTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAACGGCAAACACGCCGCCGGGTCTCCTGAGAACTCCGGCAGAGA
AAGCAAAAATAAATGCTTGACTCTGTAGCGGGAAGGCGTATTATGCACACCCCGCGCCGCT

[000348] A C1/C2 short loop on chromosome 1 whose identifier is 2218 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene clpB and has the DNA sequence

[000349] Seq. Id. = 41 Position = 1 to 72

[000350] CTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
CGGCAAACACGCCGCCGGGC

[000351] The match between the T1 sequence and the C1/C2 sequence is

[000352] Seq. Id. = 41 Position = 1 to 72

[000353] CTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
CGGCAAACACGCCGCCGGGC

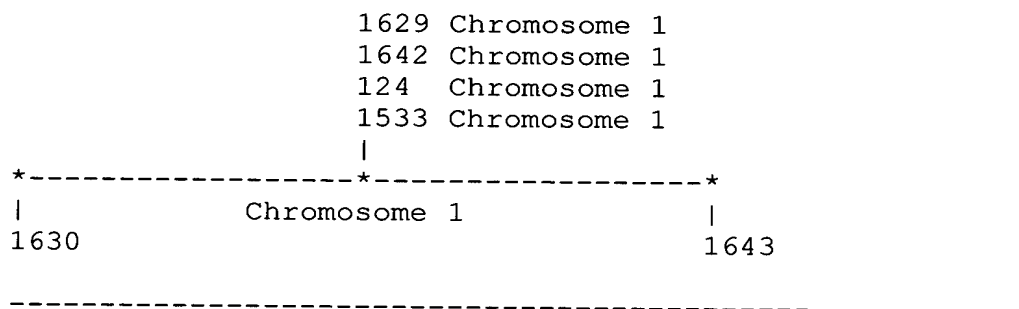
[000354] The match between the T2 sequence and the C1/C2 sequence is

[000355] Seq. Id. = 41 Position = 1 to 72

[000356] CTTGTCAGGCCGGAATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
CGGCAAACACGCCGCCGGGC

[000357] Example of a many-to-one connectron in archea - M. jannaschii

[000358] In this example the existence of the T1-T2 (1630-1643) long loop is controlled by four C1/C2 short loops (1629, 1642, 124 and 1533).



[000359] A double stranded DNA loop of length 4.998 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 1630. This T1 control element has the DNA sequence

[000360] Seq. Id. = 42 Position = 1 to 175

[000361] TTATTAATTAGTTCAAAGGATTTTATTTAATTTCTAAGGGTTTGCTGGTTT
GATTATTTAGAAATATTTGAGTTTATTGAATTATTCAGATTTTAAAAATTAAGATTAATTAG
GAAAGGAAATAAGATTTCTCTAACAGACAAGTTAAATTTTGGATTAAAAAGATAAAAAT

[000362] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1643. This T2 control element has the DNA sequence

[000363] Seq. Id. = 43 Position = 1 to 175

[000364] TTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAAATATTTGAGTTTATTG
AATTATTCAGATTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACAAA
TAAGTTAAATTTTGGATTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGAT

[000365] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

MJ1597	MJ1598	MJ1599	MJ1600	MJ1601
MJ1602				

[000366] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000367] A C1/C2 short loop on chromosome 1 whose identifier is 1629 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1597 and has the DNA sequence

[000368] Seq. Id. = 44 Position = 1 to 139

[000369] ATATGTTTGAAATTTGAAAATAAGAGTATTTAGAAGTTATTAATTAGTTCAA
AGGATTTTTATTTAATTTCTAAGGGTTTGCTGGTTTGATTATTTAGAATATTTGAGTTTATT
GAATTATTCAGATTTTTTAAAAATTA

[000370] The match between the T1 sequence and the C1/C2 sequence is

[000371] Seq. Id. = 44 Position = 37 to 139

[000372] TTATTAATTAGTTCAAAGGATTTTTATTTAATTTCTAAGGGTTTGCTGGTTT
GATTATTTAGAATATTTGAGTTTATTGAATTATTCAGATTTTTTAAAAATTA

[000373] The match between the T2 sequence and the C1/C2 sequence is

[000374] Seq. Id. = 44 Position = 81 to 139

[000375] GCTGGTTTGATTATTTAGAATATTTGAGTTTATTGAATTATTCAGATTTTAA
AAAATTA

[000376] A C1/C2 short loop on chromosome 1 whose identifier is 1642 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1602 and has the DNA sequence

[000377] Seq. Id. = 45 Position = 1 to 177

[000378] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACA
AATAAGTTAAATTTTTGGATTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGA
T

[000379] The match between the T1 sequence and the C1/C2 sequence is

[000380] Seq. Id. = 45 Position = 20 to 78

[000381] GCTGGTTTGATTATTTAGAATATTTGAGTTTATTGAATTATTCAGATTTTAA
AAAATTA

[000382] The match between the T2 sequence and the C1/C2 sequence is

[000383] Seq. Id. = 45 Position = 3 to 177

[000384] TTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTATTG
AATTATTCAGATTTTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACAAA
TAAGTTAAATTTTTGGATTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGAT

[000385] A C1/C2 short loop on chromosome 1 whose identifier is 124 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ0112 and has the DNA sequence

[000386] Seq. Id. = 46 Position = 1 to 75

[000387] ATTTAATTTCTAAGGGTTTGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTAAAT

[000388] The match between the T1 sequence and the C1/C2 sequence is

[000389] Seq. Id. = 46 Position = 1 to 75

[000390] ATTTAATTTCTAAGGGTTTGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTAAAT

[000391] The match between the T2 sequence and the C1/C2 sequence is

[000392] Seq. Id. = 46 Position = 20 to 75

[000393] GCTGGTTTGATTATTTAGAATATTTGAGTTTATTGAATTATTCAGATTTTAA
AAAT

[000394] A C1/C2 short loop on chromosome 1 whose identifier is 1533 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1486 and has the DNA sequence

[000395] Seq. Id. = 47 Position = 1 to 58

[000396] TTTTATTTAATTTCTAAGGGTTTGCTGGTTTGATTATTTAGAATATTTGAG
TTTATT

[000397] The match between the T1 sequence and the C1/C2
sequence is

[000398] Seq. Id. = 47 Position = 1 to 58

[000399] TTTTATTTAATTTCTAAGGGTTTGCTGGTTTGATTATTTAGAATATTTGAG
TTTATT

[000400] The match between the T2 sequence and the C1/C2
sequence is

[000401] Seq. Id. = 47 Position = 25 to 58

[000402] GCTGGTTTGATTATTTAGAATATTTGAGTTTATT

[000403] Example of a many-to-one connectron in single-cell
eukaryotes - *S. cerevisiae*

[000404] In this example the existence of the T1-T2 (5515-
5533) long loop on chromosome 12 is controlled by seventeen
C1/C2 short loops (5516, 5532, 1939, 2323, 1942, 3286, 3649,
4764, 4751, 5536, 6102, 8023, 7356, 3293, 3291, 3289 and 146).

5516 Chromosome 12
5532 Chromosome 12
1939 Chromosome 4
2323 Chromosome 5
1942 Chromosome 5
3286 Chromosome 7
3649 Chromosome 8

4764 Chromosome 12
 4751 Chromosome 12
 5536 Chromosome 13
 6102 Chromosome 14
 8023 Chromosome 16
 7356 Chromosome 16
 3293 Chromosome 8
 3291 Chromosome 8
 3289 Chromosome 8
 146 Chromosome 2

|
 ----------*
 | Chromosome 12 |
 3197 3308

[000405] A double stranded DNA loop of length 6.466 kilobases on chromosome 12 is bounded on the left by a T1 sequence whose identifier is 5515. This T1 control element has the DNA sequence

[000406] Seq. Id. = 48 Position = 1 to 225

[000407] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
 TTGTAAGAAATTTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
 ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
 ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000408] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 5533. This T2 control element has the DNA sequence

[000409] Seq. Id. = 49 Position = 1 to 225

[000410] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTTCTAGGGAA
 TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGC
 GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
 GG TAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000411] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

YLR467W

[000412] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000413] A C1/C2 short loop on chromosome 12 whose identifier is 5516 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YLR464W and has the DNA sequence

[000414] Seq. Id. = 50 Position = 1 to 252

[000415] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000416] A C1/C2 short loop on chromosome 12 whose identifier is 5532 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YLR467W and has the DNA sequence

[000417] Seq. Id. = 51 Position = 1 to 252

[000418] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA

ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000419] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000420] A C1/C2 short loop on chromosome 4 whose identifier is 1939 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YDR545W and has the DNA sequence

[000421] Seq. Id. = 52 Position = 1 to 222

[000422] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTGCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGG

[000423] The match between the T1 sequence and the C1/C2 sequence is

[000424] Seq. Id. = 52 Position = 1 to 222

[000425] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTGCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGG

[000426] The match between the T2 sequence and the C1/C2 sequence is

[000427] Seq. Id. = 52 Position = 28 to 222

[000428] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGG

[000429] A C1/C2 short loop on chromosome 5 whose identifier
is 2323 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YER189W and has the DNA
sequence

[000430] Seq. Id. = 53 Position = 1 to 252

[000431] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000432] The match between the T1 sequence and the C1/C2
sequence is

[000433] Seq. Id. = 53 Position = 1 to 225

[000434] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGTTG

[000435] The match between the T2 sequence and the C1/C2
sequence is

[000436] Seq. Id. = 53 Position = 28 to 252

[000437] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000438] A C1/C2 short loop on chromosome 5 whose identifier
is 1942 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YEL077C and has the DNA
sequence

[000439] Seq. Id. = 54 Position = 1 to 252

[000440] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000441] The match between the T1 sequence and the C1/C2
sequence is

[000442] Seq. Id. = 54 Position = 1 to 225

[000443] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000444] The match between the T2 sequence and the C1/C2
sequence is

[000445] Seq. Id. = 54 Position = 28 to 252

[000446] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAAGAGACAACAGGGCT

[000447] A C1/C2 short loop on chromosome 7 whose identifier
is 3286 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YGR296W and has the DNA
sequence

[000448] Seq. Id. = 55 Position = 1 to 252

[000449] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000450] The match between the T1 sequence and the C1/C2
sequence is

[000451] Seq. Id. = 55 Position = 1 to 225

[000452] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000453] The match between the T2 sequence and the C1/C2
sequence is

[000454] Seq. Id. = 55 Position = 28 to 252

[000455] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAAGAGACAACAGGGCT

[000456] A C1/C2 short loop on chromosome 8 whose identifier
is 3649 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YHR219W and has the DNA
sequence

[000457] Seq. Id. = 56 Position = 1 to 252

[000458] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000459] The match between the T1 sequence and the C1/C2
sequence is

[000460] Seq. Id. = 56 Position = 1 to 225

[000461] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000462] The match between the T2 sequence and the C1/C2
sequence is

[000463] Seq. Id. = 56 Position = 28 to 252

[000464] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000465] A C1/C2 short loop on chromosome 12 whose identifier
is 4764 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YLL066C and has the DNA
sequence

[000466] Seq. Id. = 57 Position = 1 to 252

[000467] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000468] The match between the T1 sequence and the C1/C2
sequence is

[000469] Seq. Id. = 57 Position = 1 to 225

[000470] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000471] The match between the T2 sequence and the C1/C2
sequence is

[000472] Seq. Id. = 57 Position = 28 to 252

[000473] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000474] A C1/C2 short loop on chromosome 12 whose identifier
is 4751 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YLL067C and has the DNA
sequence

[000475] Seq. Id. = 58 Position = 1 to 252

[000476] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000477] The match between the T1 sequence and the C1/C2
sequence is

[000478] Seq. Id. = 58 Position = 1 to 225

[000479] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000480] The match between the T2 sequence and the C1/C2
sequence is

[000481] Seq. Id. = 58 Position = 28 to 252

[000482] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTGTGTTGCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000483] A C1/C2 short loop on chromosome 13 whose identifier
is 5536 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YML133C and has the DNA
sequence

[000484] Seq. Id. = 59 Position = 1 to 252

[000485] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTGTGTTGCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000486] The match between the T1 sequence and the C1/C2
sequence is

[000487] Seq. Id. = 59 Position = 1 to 252

[000488] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTGTGTTGCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000489] The match between the T2 sequence and the C1/C2
sequence is

[000490] Seq. Id. = 59 Position = 28 to 252

[000491] TATAGTATATTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTA
GTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAA
AGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTT
TGGTTGAACATCCGGGTAAAGAGACAACAGGGCT

[000492] A C1/C2 short loop on chromosome 14 whose identifier
is 6102 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YNL339C and has the DNA
sequence

[000493] Seq. Id. = 60 Position = 1 to 252

[000494] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000495] The match between the T1 sequence and the C1/C2
sequence is

[000496] Seq. Id. = 60 Position = 1 to 225

[000497] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000498] The match between the T2 sequence and the C1/C2
sequence is

[000499] Seq. Id. = 60 Position = 28 to 252

[000500] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000501] A C1/C2 short loop on chromosome 16 whose identifier
is 8023 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YPR204W and has the DNA
sequence

[000502] Seq. Id. = 61 Position = 1 to 252

[000503] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000504] The match between the T1 sequence and the C1/C2
sequence is

[000505] Seq. Id. = 61 Position = 1 to 252

[000506] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000507] The match between the T2 sequence and the C1/C2
sequence is

[000508] Seq. Id. = 61 Position = 28 to 252

[000509] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000510] A C1/C2 short loop on chromosome 16 whose identifier
is 7356 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YPL283C and has the DNA
sequence

[000511] Seq. Id. = 62 Position = 1 to 252

[000512] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[000513] The match between the T1 sequence and the C1/C2
sequence is

[000514] Seq. Id. = 62 Position = 1 to 225

[000515] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000516] The match between the T2 sequence and the C1/C2
sequence is

[000517] Seq. Id. = 62 Position = 28 to 252

[000518] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTAGTGTTTGTGTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[000519] A C1/C2 short loop on chromosome 8 whose identifier
is 3293 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YHL050C and has the DNA
sequence

[000520] Seq. Id. = 63 Position = 1 to 89

[000521] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTT

[000522] The match between the T1 sequence and the C1/C2
sequence is

[000523] Seq. Id. = 63 Position = 1 to 89

[000524] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTT

[000525] The match between the T2 sequence and the C1/C2
sequence is

[000526] Seq. Id. = 63 Position = 28 to 89

[000527] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTT

[000528] A C1/C2 short loop on chromosome 8 whose identifier
is 3291 controls the expression of the genes in this T1/T2

long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YHL050C and has the DNA sequence

[000529] Seq. Id. = 64 Position = 1 to 87

[000530] ATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGCG
AGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAA

[000531] The match between the T1 sequence and the C1/C2 sequence is

[000532] Seq. Id. = 64 Position = 1 to 87

[000533] ATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGCG
AGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAA

[000534] The match between the T2 sequence and the C1/C2 sequence is

[000535] Seq. Id. = 64 Position = 1 to 87

[000536] ATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGCG
AGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAA

[000537] A C1/C2 short loop on chromosome 2 whose identifier is 145 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YBL113C and has the DNA sequence

[000538] Seq. Id. = 65 Position = 1 to 73

[000539] CTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACAT
CCGGGTAAGAGACAACAGGCT

[000540] The match between the T1 sequence and the C1/C2
sequence is

[000541] Seq. Id. = 65 Position = 1 to 47

[000542] CTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000543] The match between the T2 sequence and the C1/C2
sequence is

[000544] Seq. Id. = 65 Position = 1 to 73

[000545] CTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACAT
CCGGGTAAGAGACAACAGGCT

[000546] A C1/C2 short loop on chromosome 8 whose identifier
is 3289 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YHL050C and has the DNA
sequence

[000547] Seq. Id. = 66 Position = 1 to 73

[000548] CTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACAT
CCGGGTAAGAGACAACAGGCT

[000549] The match between the T1 sequence and the C1/C2
sequence is

[000550] Seq. Id. = 66 Position = 1 to 47

[000551] CTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[000552] The match between the T2 sequence and the C1/C2 sequence is

[000553] Seq. Id. = 66 Position = 1 to 73

[000554] CTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACAT
CCGGGTAAGAGACAACAGGCT

[000555] A C1/C2 short loop on chromosome 2 whose identifier is 146 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YBL113C and has the DNA sequence

[000556] Seq. Id. = 67 Position = 1 to 62

[000557] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAA

[000558] The match between the T1 sequence and the C1/C2 sequence is

[000559] Seq. Id. = 67 Position = 1 to 62

[000560] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAA

[000561] The match between the T2 sequence and the C1/C2 sequence is

[000562] Seq. Id. = 67 Position = 28 to 62

[000563] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAA

[000564] Example of a many-to-one connectron in multi-cell eukaryotes - C. elegans

[000565] In this example the existence of the T1-T2 (3197-3308) long loop on chromosome 5 is controlled by three C1/C2 short loops (4382, 4375 and 28633).

```

                4382  Chromosome 1
                4375  Chromosome 1
                28633 Chromosome 5
                |
*-----*-----*
|           Chromosome 5           |
28632                                     28697

```

[000566] A double stranded DNA loop of length 58.451 kilobases on chromosome 5 is bounded on the left by a T1 sequence whose identifier is 28632. This T1 control element has the DNA sequence

[000567] Seq. Id. = 68 Position = 1 to 86

[000568] GCAAAAATTGACTGAAAATTTGAATTTCCCGCAAAAATTGACTGAAAATTTGAATTTCCCGCCAAAATTGACTGAAAATTTGAA

[000569] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 28697. This T2 control element has the DNA sequence

[000570] Seq. Id. = 69 Position = 1 to 160

[000571] CAAAAAATTGACTGAAAATTTGAATTTCCCTCCAAAAATTGACTGAAAATTT
GAATTTCCCGCCAAAAATTGACTGAAAATTTGAATATCCCGCCAAAAATTGACTGAAAATTT
GAATTTCCCGCCGAAAAATTAAATGAAAAATGGAATTTCTCGCCGAA

[000572] This long T1/T2 double stranded DNA loop modulates
the expression of the following genes

M162.8	M162.4	M162.3	M162.6	M162.2
M162.1	M162.7			

[000573] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[000574] A C1/C2 short loop on chromosome 1 whose identifier
is 4382 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene Y43F8B.10 and has the DNA
sequence

[000575] Seq. Id. = 70 Position = 1 to 319

[000576] ATTATAGAAAATTTAAATTTCCCTCCAAAAATTGACTGAAAATTTGAATTT
CCCTCCAAAAATTGACTGAAAATTTGAATTTCCCGCCAAAAATTGACTGAAAATTTGAATAT
CCCGCCAAAAATTGACTGAAAATTTGAATTTCCCGCCGAAAAATTAAATGAAAAATGGAATTT
CTCGCCGAAAAATTCAGTAAAAATTTGAATTTCTTGCCAAAAATTGACTGAAAATTTGAATT
TCTTGCCAAAAAGTGACTGGGAATTTGAATTTCCCTCCAAAAATTGACTGAAAATTTGAAT
TTCCCGCTAAAAGTTGACT

[000577] The match between the T1 sequence and the C1/C2
sequence is

[000578] Seq. Id. = 70 Position = 58 to 88

[000579] CAAAAATTGACTGAAAATTTGAATTTCCCGC

[000580] The match between the T2 sequence and the C1/C2 sequence is

[000581] Seq. Id. = 70 Position = 26 to 185

[000582] CAAAAAATTGACTGAAAATTTGAATTTCCCTCCAAAAATTGACTGAAAATTT
GAATTTCCCGCCAAAAATTGACTGAAAATTTGAATATCCCGCCAAAAATTGACTGAAAATTT
GAATTTCCCGCCGAAAATTAAATGAAAAATGGAATTTCTCGCCGAA

[000583] A C1/C2 short loop on chromosome 1 whose identifier is 4375 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene Y43F8B.10 and has the DNA sequence

[000584] Seq. Id. = 71 Position = 1 to 319

[000585] ATTATAGAAAATTTAAATTTCCCTCCAAAAAATTGACTGAAAATTTGAATTT
CCCTCCAAAAAATTGACTGAAAATTTGAATTTCCCGCCAAAAAATTGACTGAAAATTTGAATAT
CCCGCCAAAAAATTGACTGAAAATTTGAATTTCCCGCCGAAAATTAAATGAAAAATGGAATTT
CTCGCCGAAAAATTCAGTAAAAATTTGAATTTCCCTGCCAAAAAATTGACTGAAAATTTGAATT
TCTTGCCAAAAAAGTGACTGGGAATTTGAATTTCCCTCCAAAAAATTGACTGAAATTTTGAAT
TTCCCGCTAAAAGTTGACT

[000586] The match between the T1 sequence and the C1/C2 sequence is

[000587] Seq. Id. = 71 Position = 58 to 88

[000588] CAAAATTGACTGAAAATTTGAATTTCCCGC

[000589] The match between the T2 sequence and the C1/C2 sequence is

[000590] Seq. Id. = 71 Position = 58 to 217

[000591] CAAAAAATTGACTGAAAATTTGAATTTCCCTCCAAAAATTGACTGAAAATTT
GAATTTCCCGCCAAAAATTGACTGAAAATTTGAATATCCCGCCAAAAATTGACTGAAAATTT
GAATTTCCCGCCGAAAATTAAATGAAAAATGGAATTTCTCGCCGAA

[000592] A C1/C2 short loop on chromosome 5 whose identifier is 28633 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene M162.5 and has the DNA sequence

[000593] Seq. Id. = 72 Position = 1 to 85

[000594] CAAAAATTGACTGAAAATTTGAATTTCCCGCAAAAAATTGACTGAAAATTTG
AATTTCCCGCCAAAAATTGACTGAAAATTTGAA

[000595] Seq. Id. = 72 Position = 1 to 85

[000596] The match between the T1 sequence and the C1/C2 sequence is

[000597] CAAAAATTGACTGAAAATTTGAATTTCCCGCAAAAAATTGACTGAAAATTTG
AATTTCCCGCCAAAAATTGACTGAAAATTTGAA

[000598] The match between the T2 sequence and the C1/C2 sequence is

[000599] Seq. Id. = 72 Position = 31 to 60

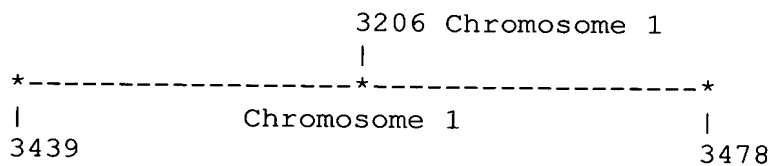
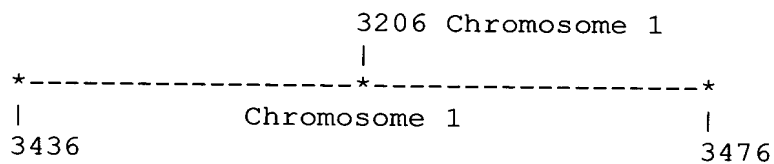
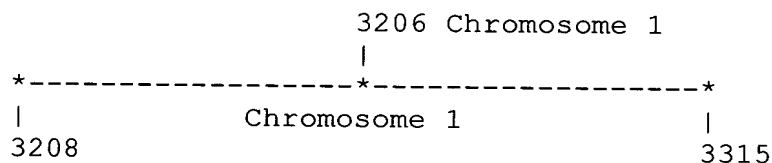
[000600] CAAAAAATTGACTGAAAATTTGAATTTCCC

3. One connectron controls the expression of many sets of genes in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[000601] One C1/C2 short loop can control the existence of a many T1-T2 long loops. The C1/C2 short loop can be on the same chromosome or on different chromosomes from the T1-T2 long loops. This relationship is described as "one-to-many". This relationship exists in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[000602] Example of a one-to-many connectron in prokaryotes - E. coli

[000603] In this example the existence of T1-T2 (3208-3315, 3436-3476, 3439-3478 and 3441-3479) long loops are controlled by one C1/C2 short loop (3206).



[000610] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

rrlC	rrfC	aspT	trpT	yifA
yifE	yifB	ilvL	ilvG_1	ilvM
ilvE	ilvD	ilvA	ilvY	ilvC
ppiC	b3776	rep	gppA	rh1B
trxA	rhoL	rho	rfe	wzzE
wecB	rffH	wecD	wecE	wzxE
yifM_2	wecG	yifK	argX	hisR
leuT	proM	aslB	aslA	hemY
hemX	hemD	cyaA	cyaY	b3808
dapF	uvrD	b3814	corA	yigF
yigG	rarD	yigI	p1dA	recQ
yigJ	yigK	p1dB	yigL	yigM
metR	metE	ysgA	udp	yigN
ubiE	yigP	b3836	yigU	yigW_1
rfaH	yigC	ubiB	fadA	fadB
pepQ	trkH	hemG	rrsA	ileT

[000611] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000612] A C1/C2 short loop on chromosome 1 whose identifier is 3206 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrsC and has the DNA sequence

[000613] Seq. Id. = 75 Position = 1 to 367

[000614] GTCCCCTTCGTCTAGAGGCCACAGGACACCGCCCTTTCACGGCGGTAACAGGG
 GTTCGAATCCCCTAGGGGACGCCACTTGCTGGTTTGTGAGTGAAAGTCACCTGCCTTAATAT
 CTCAAAACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAAGCTG
 AAAATTGAAA...ACGGCGATTTCCGAATGGGGAAACCCAGTGTGTTTCGACACACTATCA

TTAACTGAATCCATAGGTTAATGAGGCGAACCGGGGGAAGTGAACATCTAAGTACCCCGAG
GAAAAGAAATCAACCGAGATTCCCCCAGTAGCGGCGAGCGAACGGGGAGCAGCCCAGAGCCT
GAATCAGT

[000615] The match between the T1 sequence and the C1/C2
sequence is

[000616] Seq. Id. = 75 Position = 121 to 367

[000617] ACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAA
GCTGAAAATTGAAACACTGAACAACGAAAGTTGTTTCGTGAGTCTCTCAAATTTTCGCAACAC
GATGATGAATCGAAAGAAACATCTTCGGGTTGTGAGGTTAAGCGACTAAGCGTACACGGTGG
ATGCCCTGGC...AGTGTGTTTCGACACACTATCATTAAGTGAATCCATAGGTTAATGAGGC
GAACCGGGGGAAGTGAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCCCCA
GTAGCGGCGAGCGAACGGGGAGCAGCCCAGAGCCTGAATCAGT

[000618] The match between the T2 sequence and the C1/C2
sequence is

[000619] Seq. Id. = 75 Position = 148 to 232

[000620] TTTGCTCTTTAAAAATCTGGATCAAGCTGAAAATTGAAACACTGAACAACGA
AAGTTGTTTCGTGAGTCTCTCAAATTTTCGCAAC

[000621] A double stranded DNA loop of length 41.279 kilo-
bases on chromosome 1 is bounded on the left by a T1 sequence
whose identifier is 3436. This T1 control element has the DNA
sequence

[000622] Seq. Id. = 76 Position = 1 to 113

[000623] ACGCAACGCGTGATAAGCAATTTTCGTGTCCCCTTCGTCTAGAGGCCAGGA
CACCGCCCTTTACGGCGGTAACAGGGGTTTCGAATCCCCTAGGGGACGCCACTTGCTGGTT

[000624] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3476. This T2 control element has the DNA sequence

[000625] Seq. Id. = 77 Position = 1 to 150

[000626] AGTGAAAAGCAAGGCGTCTTGCGAAGCAGACTGATACGTCCCCTTCGTCTAG
AGGCCAGGACACCGCCCTTTACGGCGGTAACAGGGGTTTCGAATCCCCTAGGGGACGCCAC
TTGCTGGTTTGTGAGTGAAAGTCACCTGCCTTAATA

[000627] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

gltT	rrlB	rrfB	murB	coaA
b3975	tyrU	thrT	tufB	secE
nusG	rplK	rplA	rplJ	rplL
rpoB	rpoC	htrC	thiH	thiF
thiE	yjaE	yjaD	hemE	nfi
yjaG	hupA	yjaH	yjaI	hydH
purD	purH			

[000628] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000629] A C1/C2 short loop on chromosome 1 whose identifier is 3206 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrsC and has the DNA sequence

[000630] Seq. Id. = 78 Position = 1 to 553

[000631] GTCCCCTTCGTCTAGAGGCCAGGACACCGCCCTTTCACGGCGGTAACAGGG
GTTTGAATCCCCTAGGGGACGCCACTTGCTGGTTTGTGAGTGAAAGTCACCTGCCTTAATAT
CTCAAACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAAGCTG
AAAATTGAAACACTGAACAACGAAAGTTGTTTCGTGAGTCTCTCAAATTTTCGCAACACGATG
ATGAATCGAAAGAAACATCTTCGGGTGAGGTTAAGCGACTAAGCGTACACGGTGGATGC
CCTGGCAGTCAGAGGCGATGAAGGACGTGCTAATCTGCGATAAGCGTCGGTAAGGTGATATG
AACCGTTATAACCGGCGATTTCCGAATGGGGAAACCCAGTGTGTTTCGACACACTATCATTA
ACTGAATCCATAGGTAAATGAGGCGAACCGGGGAACTGAAACATCTAAGTACCCCGAGGAA
AAGAAATCAACCGAGATTCCCCCAGTAGCGGCGAGCGAACGGGGAGCAGCCCAGAGCCTGAA
TCAGT

[000632] The match between the T1 sequence and the C1/C2
sequence is

[000633] Seq. Id. = 78 Position = 1 to 86

[000634] GTCCCCTTCGTCTAGAGGCCAGGACACCGCCCTTTCACGGCGGTAACAGGG
GTTTGAATCCCCTAGGGGACGCCACTTGCTGGTT

[000635] The match between the T2 sequence and the C1/C2
sequence is

[000636] Seq. Id. = 78 Position = 1 to 113

[000637] GTCCCCTTCGTCTAGAGGCCAGGACACCGCCCTTTCACGGCGGTAACAGGG
GTTTGAATCCCCTAGGGGACGCCACTTGCTGGTTTGTGAGTGAAAGTCACCTGCCTTAATA

[000638] A double stranded DNA loop of length 41.336 kilo-
bases on chromosome 1 is bounded on the left by a T1 sequence
whose identifier is 3439. This T1 control element has the DNA
sequence

[000639] Seq. Id. = 79 Position = 1 to 94

[000640] CCTTAATATCTCAAACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTT
AAAAATCTGGATCAAGCTGAAAATTGAAACACTGAACAACGA

[000641] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3478. This T2 control element has the DNA sequence

[000642] Seq. Id. = 80 Position = 1 to 94

[000643] GTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAAGCTGAAAATTGA
AACACTGAACAACGAAAGTTGTTCGTGAGTCTCTCAAATTTT

[000644] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

rrlB	rrfB	murB	coaA	b3975
tyrU	thrT	tufB	secE	nusG
rplK	rplA	rplJ	rplL	rpoB
rpoC	htrC	thiH	thiF	thiE
yjaE	yjaD	hemE	nfi	yjaG
hupA	yjaH	yjaI	hydH	purD
purH	gltV			

[000645] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000646] A C1/C2 short loop on chromosome 1 whose identifier is 3206 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the generrc and has the DNA sequence

[000647] Seq. Id. = 81 Position = 1 to 367

[000648] GTCCCCTTCGTCTAGAGGCCAGGACACCGCCCTTTCACGGCGGTAACAGGG
GTTTCGAATCCCCTAGGGGACGCCACTTGCTGGTTTGTGAGTGAAAGTCACCTGCCTTAATAT
CTCAAACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAAGCTG
AAAATTGAAA...ACCGGCGATTTCCGAATGGGGAAACCCAGTGTGTTTCGACACACTATCA
TTAACTGAATCCATAGGTTAATGAGGCGAACCGGGGGAACCTGAAACATCTAAGTACCCCGAG
GAAAAGAAATCAACCGAGATCCCCCAGTAGCGGCGAGCGAACGGGGAGCAGCCCAGAGCCT
GAATCAGT

[000649] The match between the T1 sequence and the C1/C2
sequence is

[000650] Seq. Id. = 81 Position = 106 to 199

[000651] CCTTAATATCTCAAACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTT
AAAAATCTGGATCAAGCTGAAAATTGAAACACTGAACAACGA

[000652] The match between the T2 sequence and the C1/C2
sequence is

[000653] Seq. Id. = 81 Position = 133 to 226

[000654] GTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAAGCTGAAAATTGA
AACTGAACAACGAAAGTTGTTTCGTGAGTCTCTCAAATTTT

[000655] A double stranded DNA loop of length 38.285 kilo-
bases on chromosome 1 is bounded on the left by a T1 sequence
whose identifier is 3441. This T1 control element has the DNA
sequence

[000656] Seq. Id. = 82 Position = 1 to 355

[000657] AATTTTCGCAACACGATGATGAATCGAAAGAAACATCTTCGGGTGTGAGGT
 TAAGCGACTAAGCGTACACGGTGGATGCCCTGGCAGTCAGAGGCGATGAAGGACGTGCTAAT
 CTGCGATAAGCGTCGGTAAGGTGATATGAACCGTTATAACCGGCGATTTCCGAATGGGGAAA
 CCCAGTGTGT...GATGAGAGAAGATTTTCAGCCTGATACAGATTAAATCAGAACGCAGAAG
 CGGTCTGATAAAACAGAATTTGCCTGGCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGC
 CGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAG

[000658] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3479. This T2 control element has the DNA sequence

[000659] Seq. Id. = 83 Position = 1 to 356

[000660] AAGAAACATCTTCGGGTGTGAGGTTAAGCGACTAAGCGTACACGGTGGATG
 CCCTGGCAGTCAGAGGCGATGAAGGACGTGCTAATCTGCGATAAGCGTCGGTAAGGTGATAT
 GAACCGTTATAACCGGCGATTTCCGAATGGGGAAACCCAGTGTGTTTCGACACACTATCATT
 AACTGAATCC...CAGATTAAATCAGAACGCAGAAGCGGTCTGATAAAACAGAATTTGCCTG
 GCGGCAGTAGCGCGGTGGTCCCACCTGACCCCATGCCGAACCTCAGAAGTGAAACGCCGTAGC
 GCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACCTGCCAGGCATCAAATTA

[000661] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

rrlB	rrfB	murB	coaA	b3975
tyrU	thrT	tufB	secE	nusG
rplK	rplA	rplJ	rplL	rpoB
rpoC	htrC	thiH	thiF	thiE
yjaE	yjaD	hemE	nfi	yjaG
hupA	yjaH	yjaI	hydH	purD
purH	gltV			

[000662] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000663] A C1/C2 short loop on chromosome 1 whose identifier is 3206 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrsC and has the DNA sequence

[000664] Seq. Id. = 84 Position = 1 to 519

[000665] GTCCCCTTCGTCTAGAGGCCAGGACACCGCCCTTTCACGGCGGTAACAGGG
GTTCGAATCCCCTAGGGGACGCCACTTGCTGGTTTGTGAGTGAAAGTCACCTGCCTTAATAT
CTCAAACTCATCTTCGGGTGATGTTTGAGATATTTGCTCTTTAAAAATCTGGATCAAGCTG
AAAATTGAAAAATTTTCGCAACACGATGATGAATCGAAAGAAACATCTTCGGGTGTGAGGT
TAAGCGACTAAGCGTACACGGTGGATGCCCTGGCAGTCAGAGGCGATGAAGGACGTGCTAAT
CTGCGATAAGCGTCGGTAAGGTGATATGAACCGTTATAACCGGCGATTTCCGAATGGGGAAA
CCCAGTGTGTTTCGACACACTATCATTAAGTGAATCCATAGGTTAATGAGGCGAACCGGGGG
AACTGAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCCCCAGTAGCGGCGA
GCGAACGGGGAGCAGCCCAGAGCCTGAATCAGT

[000666] The match between the T1 sequence and the C1/C2 sequence is

[000667] Seq. Id. = 84 Position = 187 to 519

[000668] AATTTTCGCAACACGATGATGAATCGAAAGAAACATCTTCGGGTGTGAGGT
TAAGCGACTAAGCGTACACGGTGGATGCCCTGGCAGTCAGAGGCGATGAAGGACGTGCTAAT
CTGCGATAAGCGTCGGTAAGGTGATATGAACCGTTATAACCGGCGATTTCCGAATGGGGAAA
CCCAGTGTGTTTCGACACACTATCATTAAGTGAATCCATAGGTTAATGAGGCGAACCGGGGG
AACTGAAACATCTAAGTACCCCGAGGAAAAGAAATCAACCGAGATTCCCCCAGTAGCGGCGA
GCGAACGGGGAGCAGCCCAGAGCCTGAATCAGT

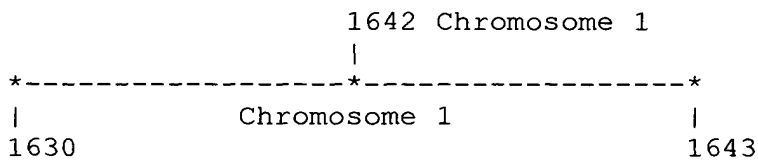
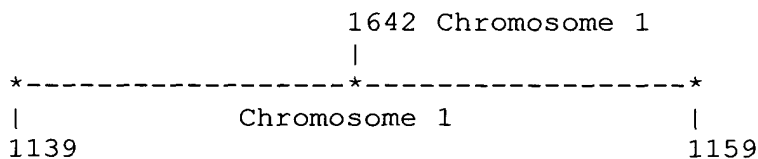
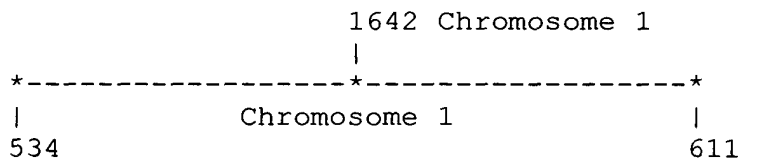
[000669] The match between the T2 sequence and the C1/C2 sequence is

[000670] Seq. Id. = 84 Position = 214 to 519

[000671] AAGAAACATCTTCGGGTTGTGAGGTTAAGCGACTAAGCGTACACGGTGGATG
 CCCTGGCAGTCAGAGGCGATGAAGGACGTGCTAATCTGCGATAAGCGTCGGTAAGGTGATAT
 GAACCGTTATAACCGGCGATTTCCGAATGGGGAAACCCAGTGTGTTTCGACACACTATCATT
 AACTGAATCCATAGGTTAATGAGGCGAACCGGGGGAACCTGAAACATCTAAGTACCCCGAGGA
 AAAGAAATCAACCGAGATTCCCCCAGTAGCGGCGAGCGAACGGGGAGCAGCCCAGAGCCTGA
 ATCAGT

[000672] Example of a one-to-many connectron in archea - M.
 jannaschii

[000673] In this example the existence of T1-T2 (534-611,
 1139-1159, and 1630-1643) long loops are controlled by one
 C1/C2 short loop (1642).



[000674] A double stranded DNA loop of length 72.886 kilo-
 bases on chromosome 1 is bounded on the left by a T1 sequence

whose identifier is 534. This T1 control element has the DNA sequence

[000675] Seq. Id. = 85 Position = 1 to 37

[000676] TAAGTAAATAAAATTTCTCTAACAAATAAGTTAAATT

[000677] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 611. This T2 control element has the DNA sequence

[000678] Seq. Id. = 86 Position = 1 to 59

[000679] TAAATAAAATTTCTCTAACAAATAAGTTAAATTTTGGATTAAAAAGATAA
AAATGCT

[000680] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

MJ0486	MJ0487	MJ0488	MJ0489	MJ0490
MJ0492	MJ0493	MJ0494	MJ0495	MJ0496
MJ0497	MJ0499	MJ0500	MJ0501	MJ0502
MJ0503	MJ0504	MJ0506	MJ0507	MJ0508
MJ0509	MJ0510	MJ0511	MJ0512	MJ0513
MJ0514	MJ0514	MJ0517	MJ0519	MJ0520
MJ0521	MJ0522	MJ0523	MJ0525	MJ0526
MJ0526	MJ0529	MJ0530	MJ0531	MJ0532
MJ0534	MJ0535	MJ0536	MJ0538	MJ0539
MJ0540	MJ0541	MJ0542	MJ0543	MJ0544
MJ0545	MJ0547	MJ0548	MJ0549	MJ0550
MJ0552	MJ0553	MJ0554	MJ0555	MJ0556
MJ0558	MJ0559	MJ0560	MJ0561	MJ0562
MJ0563	MJ0564			

[000681] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000682] A C1/C2 short loop on chromosome 1 whose identifier is 1642 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1602 and has the DNA sequence

[000683] Seq. Id. = 87 Position = 1 to 177

[000684] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACA
AATAAGTTAAATTTTTGGATTTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGA
T

[000685] The match between the T1 sequence and the C1/C2 sequence is

[000686] Seq. Id. = 87 Position = 92 to 127

[000687] AAGTAAATAAAATTTCTCTAACAATAAGTTAAATT

[000688] The match between the T2 sequence and the C1/C2 sequence is

[000689] Seq. Id. = 87 Position = 95 to 150

[000690] TAAATAAAATTTCTCTAACAATAAGTTAAATTTTTGGATTTAAAAAGATAA
AAAT

[000691] A double stranded DNA loop of length 14.509 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 1139. This T1 control element has the DNA sequence

[000692] Seq. Id. = 88 Position = 1 to 78

[000693] ATTTATTAATTAGTTCAAAGGATTTTTATTTAATTTCTAAGGGTTAGCTGGT
TTGATTGTTTAAAATATTTGAGTTTA

[000694] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1159. This T2 control element has the DNA sequence

[000695] Seq. Id. = 89 Position = 1 to 78

[000696] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAAATATTTGAGTTTAT
TGAATTATTCAGATTTTTTAAAAATTA

[000697] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

MJ1096	MJ1097	tRNA-Arg-3	MJ1098	MJ1099
MJ1100	MJ1101	MJ1102	MJ1103	MJ1104
MJ1105	MJ1106	MJ1107	MJ1108	

[000698] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000699] A C1/C2 short loop on chromosome 1 whose identifier is 1642 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1602 and has the DNA sequence

[000700] Seq. Id. = 90 Position = 1 to 177

[000701] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACA
AATAAGTTAAATTTTTGGATTTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGA
T

[000702] The match between the T1 sequence and the C1/C2
sequence is

[000703] Seq. Id. = 90 Position = 1 to 31

[000704] ATTTAATTTCTAAGGGTTAGCTGGTTTGATT

[000705] The match between the T2 sequence and the C1/C2
sequence is

[000706] Seq. Id. = 90 Position = 1 to 78

[000707] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTTTAAAAATTA

[000708] A double stranded DNA loop of length 4.998 kilo-
bases on chromosome 1 is bounded on the left by a T1 sequence
whose identifier is 1630. This T1 control element has the DNA
sequence

[000709] Seq. Id. = 91 Position = 1 to 175

[000710] TTATTAATTAGTTCAAAGGATTTTATTTAATTTCTAAGGGTTTGCTGGTTT
GATTATTTAGAATATTTGAGTTTATTGAATTATTCAGATTTTAAAAATTAAGATTAATTAG
GAAAGGAAATAAGATTTCTCTAACAGACAAGTTAAATTTTGGATTTAAAAAGATAAAAAT

[000711] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1643. This T2 control element has the DNA sequence

[000712] Seq. Id. = 92 Position = 1 to 175

[000713] TTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTATTG
AATTATTCAGATTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACAAA
TAAGTTAAATTTTGGATTTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGAT

[000714] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

MJ1597	MJ1598	MJ1599	MJ1600	MJ1601
MJ1602				

[000715] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000716] A C1/C2 short loop on chromosome 1 whose identifier is 1642 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1602 and has the DNA sequence

[000717] Seq. Id. = 93 Position = 1 to 177

[000718] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACA

AATAAGTTAAATTTTTGGATTTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGAT
T

[000719] The match between the T1 sequence and the C1/C2 sequence is

[000720] Seq. Id. = 93 Position = 20 to 78

[000721] GCTGGTTTGATTATTTAGAATATTTGAGTTTATTGAATTATTCAGATTTTAA
AAAATTA

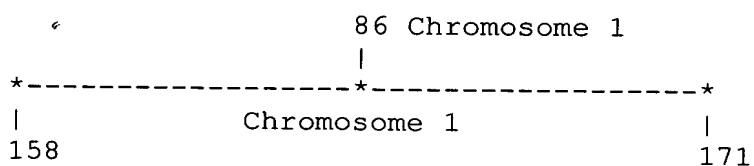
[000722] The match between the T2 sequence and the C1/C2 sequence is

[000723] Seq. Id. = 93 Position = 3 to 177

[000724] TTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTATTG
AATTATTCAGATTTTAAAAATTAGGATTAATTAGGCAAGTAAATAAAATTTCTCTAACAAA
TAAGTTAAATTTTTGGATTTAAAAAGATAAAAATACTCTGTTTTATTATGGAAAGAAAGAT

[000725] Example of a one-to-many connectron in single-cell eukaryotes - *S. cerevisiae*

[000726] In this example the existence of T1-T2 (158-171, 293-317, 4295-4308 and 5916-5923) long loops are controlled by one C1/C2 short loop (86).



86 Chromosome 1
 |

 | Chromosome 1 |
 293 317

86 Chromosome 1
 |

 | Chromosome 10 |
 4295 4308

86 Chromosome 1
 |

 | Chromosome 13 |
 5916 5923

[000727] A double stranded DNA loop of length 20.391 kilobases on chromosome 2 is bounded on the left by a T1 sequence whose identifier is 158. This T1 control element has the DNA sequence

[000728] Seq. Id. = 94 Position = 1 to 153

[000729] CCAATTGTTGGAATAAAAAATCAACTATCATCTACTAACTAGTATTTACGTTA
 CTAGTATATTATCATATACGGTGTTAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAA
 TTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAG

[000730] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 171. This T2 control element has the DNA sequence

[000731] Seq. Id. = 95 Position = 1 to 192

[000732] ATAATTGTTGGAATAAAAATCAACTATCATCTACTAACTAGTATTTACGTTA
CTAGTATATTATCATATACGGTGTGTTAGAAGATGACACAAATGATGAGAAATAGTCATCTAAA
TTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATTAACATATAA
AATGATGATAATAATA

[000733] This long T1/T2 double stranded DNA loop modulates
the expression of the following genes

YBL107W-A	TL(UAA)B1	YBL107C	YBL106C	YBL105C
YBL104C	YBL103C	YBL102W	YBL101C	

[000734] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[000735] A C1/C2 short loop on chromosome 1 whose identifier
is 86 controls the expression of the genes in this T1/T2 long
loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YAR009C and has the DNA
sequence

[000736] Seq. Id. = 96 Position = 1 to 362

[000737] ATCTATTACATTATGGGTGGTATGTTGGAATAGAAATCAACTATCATCTACT
AACTAGTATTTACATTACTAGTATATTATCATATACGGTGTGTTAGAAGATGACGCAAATGATG
AGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAA
TGAATATAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGA
TTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAATA
TTATAGCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTCACCCATTTCTCAGAA

[000738] The match between the T1 sequence and the C1/C2
sequence is

[000739] Seq. Id. = 96 Position = 34 to 65

[000740] AAATCAACTATCATCTACTAACTAGTATTTAC

[000741] The match between the T2 sequence and the C1/C2 sequence is

[000742] Seq. Id. = 96 Position = 34 to 65

[000743] AAATCAACTATCATCTACTAACTAGTATTTAC

[000744] A double stranded DNA loop of length 38.470 kilobases on chromosome 2 is bounded on the left by a T1 sequence whose identifier is 293. This T1 control element has the DNA sequence

[000745] Seq. Id. = 97 Position = 1 to 258

[000746] GAATTGTTGGAATAAAAATCCACTATCGTCTATCAACTAATAGTTATATTAT
CAATATATTATCATATACGGTGTTAAGATGATGACATAAGTTATGAGAAGCTGTCATCGAAG
TTAGAGGAAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAATGAAACATATAAAACGGA
ATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATAT
CCTTGAGGAGAACTTCTAGT

[000747] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 317. This T2 control element has the DNA sequence

[000748] Seq. Id. = 98 Position = 1 to 77

[000749] AATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTCG
AGGAGAACTTCTAGTATATTCTGTA

[000750] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

YBL005W-B	TS(AGA)B	YBL004W	YBL003C	YBL002W
YBL001C	YBR001C	YBR002C	YBR003W	YBR004C
YBR005W	YBR006W	YBR007C	YBR008C	YBR009C
YBR010W	YBR011C	YBR012C		

[000751] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000752] A C1/C2 short loop on chromosome 1 whose identifier is 86 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YAR009C and has the DNA sequence

[000753] Seq. Id. = 99 Position = 1 to 362

[000754] ATCTATTACATTATGGGTGGTATGTTGGAATAGAAATCAACTATCATCTACT
AACTAGTATTTACATTACTAGTATATTATCATATACGGTGTTAGAAGATGACGCAAATGATG
AGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAA
TGAATATAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGA
TTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAATA
TTATAGCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTCACCCATTTCTCAGAA

[000755] The match between the T1 sequence and the C1/C2 sequence is

[000756] Seq. Id. = 99 Position = 181 to 264

[000757] AAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATAT
AGATTCCATTTTGAGGATTCCTATATCCT

[000758] The match between the T2 sequence and the C1/C2 sequence is

[000759] Seq. Id. = 99 Position = 215 to 291

[000760] AATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTCG
AGGAGAACTTCTAGTATATTCTGTA

[000761] A double stranded DNA loop of length 11.020 kilobases on chromosome 10 is bounded on the left by a T1 sequence whose identifier is 4295. This T1 control element has the DNA sequence

[000762] Seq. Id. = 100 Position = 1 to 145

[000763] AAACGCAAGGATTGATAATGTAATAGGATCAATGAATATAAACATATAAAAC
GGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTA
TATCCTCGAGGAGAACTTCTAGTATATTCTG

[000764] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 4308. This T2 control element has the DNA sequence

[000765] Seq. Id. = 101 Position = 1 to 180

[000766] GGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATAAACA
TATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCATTTTGAG
GATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAATATTATAGCCTTTA
TCAA

[000767] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

YJR027W YJR029W

[000768] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000769] A C1/C2 short loop on chromosome 1 whose identifier is 87 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YAR009C and has the DNA sequence

[000770] Seq. Id. = 102 Position = 1 to 359

[000771] ATCTATTACATTATGGGTGGTATGTTGGAATAGAAATCAACTATCATCTACT
AACTAGTATTTACATTACTAGTATATTATCATATACGGTGTTAGAAGATGACGCAAATGATG
AGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAA
TGAATATAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGA
TTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAATA
TTATAGCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTCACCCATTTCTCA

[000772] A double stranded DNA loop of length 5.462 kilo-
bases on chromosome 13 is bounded on the left by a T1 sequence
whose identifier is 5916. This T1 control element has the DNA
sequence

[000773] Seq. Id. = 103 Position = 1 to 146

[000774] AAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAATGAAACATATAAAAC
GGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTA
TATCCTCGAGGAGAACTTCTAGTATATTCTGTA

[000775] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 5923. This T2 control element has the DNA sequence

[000776] Seq. Id. = 104 Position = 1 to 146

[000777] 104TAATAGGATAATGAAACATATAAAACGGAATGAGGAATAATCGTAATAT
TAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGT
ATATTCTGTATACCTAATATTATAGCCTTTATCAA

[000778] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

YML045W

[000779] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

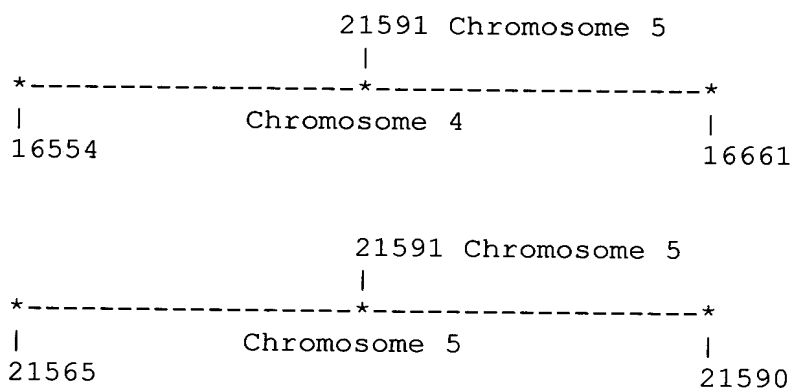
[000780] A C1/C2 short loop on chromosome 1 whose identifier is 87 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YAR009C and has the DNA sequence

[000781] Seq. Id. = 105 Position = 1 to 359

[000782] ATCTATTACATTATGGGTGGTATGTTGGAATAGAAATCAACTATCATCTACT
AACTAGTATTTACATTACTAGTATATTATCATATACGGTGTTAGAAGATGACGCAAATGATG
AGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAA
TGAATATAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGA
TTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAATA
TTATAGCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTCACCCATTTCTCA

[000783] Example of a one-to-many connectron in multi-cell eukaryotes - *C. elegans*

[000784] In this example the existence of T1-T2 (16554-16661 and 21565-21590) long loops are controlled by one C1/C2 short loop (21591).



[000785] A double stranded DNA loop of length 50.159 kilobases on chromosome 4 is bounded on the left by a T1 sequence whose identifier is 16554. This T1 control element has the DNA sequence

[000786] Seq. Id. = 106 Position = 1 to 143

[000787] TGCCTGAAAAAATTGGCTCCGAGTTAGGACACTTGGGGTGGTCAAAAAATTT
TGTGACTATTGTCAAATGAAAGATCATAGTTGATAACATAAATTCCCAAAGTTTCATAAAAA
TCGATACGCAGCGAACAAAGTTATCAATT

[000788] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 16661. This T2 control element has the DNA sequence

[000789] Seq. Id. = 107 Position = 1 to 141

[000790] CACTTGGGGTGGTCAAAAAATTTTGTGATTATTGTCAAATGAAAGATCATGG
TTGATAACATAAATTCCCAAAGTTTCATAAAAATCGATACGCAGCGAACAAAGTTATGATTT
TTGACCCGGAACCTTATTTGGAGACCTA

[000791] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

C23H5.7 C23H5.8a C23H5.3 C23H5.2 C23H5.9
C23H5.1

[000792] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000793] A C1/C2 short loop on chromosome 5 whose identifier is 21591 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene F25A2.1 and has the DNA sequence

[000794] Seq. Id. = 108 Position = 1 to 117

[000795] TATTGTCAAATGAAAGATCATGGTTGATAACATAAATTCCCACAATTCATA
AAAATCGATACGCAGCGAACAAAGTTATGATTTTTGACCCGGAACCTTATTTGGAGACCTAAT
ATT

[000796] The match between the T1 sequence and the C1/C2 sequence is

[000797] Seq. Id. = 108 Position = 46 to 85

[000798] TTTCATAAAAATCGATACGCAGCGAACAAAGTTAT

[000799] The match between the T2 sequence and the C1/C2 sequence is

[000800] Seq. Id. = 108 Position = 1 to 42

[000801] TATTGTCAAATGAAAGATCATGGTTGATAACATAAATTCCCA

[000802] A double stranded DNA loop of length 18.142 kilobases on chromosome 5 is bounded on the left by a T1 sequence whose identifier is 21565. This T1 control element has the DNA sequence

[000803] Seq. Id. = 109 Position = 1 to 72

[000804] CTCCGAGTTAGGACACTTGGGGTGGACAAAAAATTTTGTGACTATTGTCAAA
TGAAAGATCATGGTTGATAA

[000805] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 21590. This T2 control element has the DNA sequence

[000806] Seq. Id. = 110 Position = 1 to 115

[000807] TATTGTCAAATGAAAGATCATGGTTGATAACATAAATTCCCACAATTTTCATA
AAAATCGATACGCAGCGAACAAAGTTATGATTTTTGACCCGGAAGTTATTTGGAGACCTAAT
A

[000808] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

T21H3.2 T21H3.1 F25A2.1

[000809] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000810] A C1/C2 short loop on chromosome 5 whose identifier is 21591 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene F25A2.1 and has the DNA sequence

[000811] Seq. Id. = 111 Position = 1 to 117

[000812] TATTGTCAAATGAAAGATCATGGTTGATAACATAAATTCCCACAATTCATA
AAAATCGATACGCAGCGAACAAAGTTATGATTTTTGACCCGGAAGTTATTTGGAGACCTAAT
ATT

[000813] The match between the T1 sequence and the C1/C2 sequence is

[000814] Seq. Id. = 111 Position = 1 to 30

[000815] TATTGTCAAATGAAAGATCATGGTTGATAA

[000816] The match between the T2 sequence and the C1/C2 sequence is

[000817] Seq. Id. = 111 Position = 1 to 115

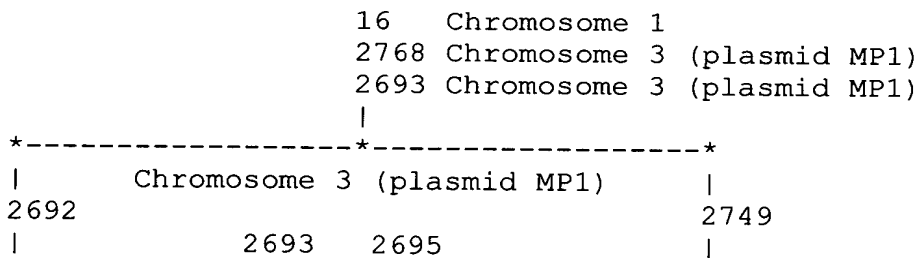
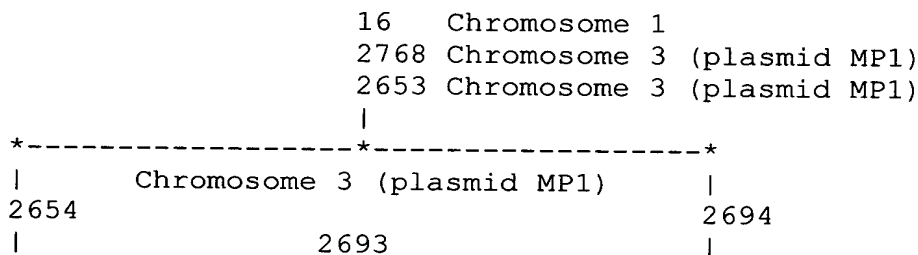
[000818] TATTGTCAAATGAAAGATCATGGTTGATAACATAAATTCCCACAATTCATA
AAAATCGATACGCAGCGAACAAAGTTATGATTTTTGACCCGGAACCTATTTGGAGACCTAAT
A

4. Connectrons occur between prokaryotes and their plasmids.

[000819] Connectron relationships exist between prokaryotes and their plasmids. These connectrons implement a control mechanism between the two genomes that makes it possible for them to form a symbiotic relationship. In the case of *D. radiodurans* the relationship is not symmetric. The *D. radiodurans* genome sends C1/C2 short loops to the MP1 plasmid.

[000820] Example of a prokaryote/plasmid connectron - *D. radiodurans*

[000821] In this example the existence of T1-T2 (2654-2694 and 2692-2749) long loops in chromosome 3 that is the plasmid MP1 are controlled by one C1/C2 short loop (16) in chromosome 1.



[000822] A double stranded DNA loop of length 46.903 kilobases on chromosome 3 (plasmid MP1) is bounded on the left by a T1 sequence whose identifier is 2654. This T1 control element has the DNA sequence

[000823] Seq. Id. = 112 Position = 1 to 274

[000824] CAGCGTTTTTCTCGCTGTTTCCTGGACGGCTGAACGCCCTGAATCTCTCCCGG
TATGCAGCCTGCTCGGAGAGTACGATTCGTCTGTTGGCTGCACCGAAGTGACGATGGGGCCAT
TCCGTGGGGCGCGTTACACCAGGCGACTGTCTAGTACAGCAATCGAGAGTGGGCTGATCAGCC
CACTGTGCGTTCTGGCCATCGACGCCTCTTTTCACCGCAAAGCCGGTCAGCACACCGCACAC
CTCGGCTCGTTCTGGAATGGCTGTGCCGCGCGGACC

[000825] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 2694. This T2 control element has the DNA sequence

[000826] Seq. Id. = 113 Position = 1 to 274

[000827] GCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAGTACGATT
CGTCGTTGGCTGCACCGAAGTGACGATGGGGCCATTCCGTGGGGCGCGTTACACCAGGCGAC
TGTCAGTACAGCAATCGAGAGTGGGCTGATCAGCCCACTGTGCGTTCTGGCCATCGACGCCT
CTTTTCACCGCAAAGCCGGTCAGCACACCGCACACCTCGGCTCGTTCTGGAATGGCTGTGCC
GCGCGGACCGAACGCGGAATCGAGCAATCCTGTTGT

[000828] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

DRB0020	DRB0021	DRB0022	DRB0023	DRB0024
DRB0025	DRB0027	DRB0030	DRB0032	DRB0033
DRB0034	DRB0035	DRB0037	DRB0038	DRB0039
DRB0041	DRB0042	DRB0043	DRB0044	DRB0045
DRB0047	DRB0051	DRB0052	DRB0054	DRB0055
DRB0057				

[000829] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000830] A C1/C2 short loop on chromosome 3 (plasmid MP1) whose identifier is 2693 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene DRB0057 and has the DNA sequence

[000831] Seq. Id. = 114 Position = 1 to 103

[000832] CTGATGGCCATCCTACAGTACGTTCTCAGCGCGGTCCCGCTGCGCAAGACGC
AGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTTCTCGCTGTTCTGAC

[000833] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000834] A C1/C2 short loop on chromosome 1 whose identifier is 16 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene DR0009 and has the DNA sequence

[000835] Seq. Id. = 115 Position = 1 to 186

[000836] GCTGTGAAATCACCGCTTCCAATGGGTCTGATGGCCATCCTACAGTACGTTT
TCAGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTT
CTCGCTGTTCTGGACGGCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAG
TACGATTCGT

[000837] The match between the T1 sequence and the C1/C2 sequence is

[000838] Seq. Id. = 115 Position = 105 to 186

[000839] CAGCGTTTTTCTCGCTGTTCCCTGGACGGCTGAACGCCCTGAATCTCTCCCGG
TATGCAGCCTGCTCGGAGAGTACGATTCGT

[000840] The match between the T2 sequence and the C1/C2
sequence is

[000841] Seq. Id. = 115 Position = 132 to 186

[000842] GCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAGTACGATT
CGT

[000843] A C1/C2 short loop on chromosome 3 (plasmid MP1)
whose identifier is 2768 controls the expression of the genes
in this T1/T2 long loop. This C1/C2 short loop is expressed
as a RNA single strand that is 3'UTR to the gene DRB0133 and
has the DNA sequence

[000844] Seq. Id. = 116 Position = 1 to 186

[000845] GCTGTGAAATCACCGCTTCCAATGGGTCTGATGGCCATCCTACAGTACGTTC
TCAGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTT
CTCGCTGTTCCCTGGACGGCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAG
TACGATTCGT

[000846] The match between the T1 sequence and the C1/C2
sequence is

[000847] Seq. Id. = 116 Position = 105 to 186

[000848] CAGCGTTTTTCTCGCTGTTCCCTGGACGGCTGAACGCCCTGAATCTCTCCCGG
TATGCAGCCTGCTCGGAGAGTACGATTCGT

[000849] The match between the T2 sequence and the C1/C2 sequence is

[000850] Seq. Id. = 116 Position = 132 to 186

[000851] GCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAGTACGATT
CGT

[000852] A C1/C2 short loop on chromosome 3 (plasmid MP1) whose identifier is 2653 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene DRB0017 and has the DNA sequence

[000853] Seq. Id. = 117 Position = 1 to 186

[000854] CGGTCCCGCTGCGCAAGACGCAGCGGAATTCCTGACCGTGCTGCTCAGCGT
TTTTCTCGCTGTTCTGGACGGCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGG
AGAGTACGATTCGTCGTTGGCTGCACCGAAGTGACGATGGGGCCATTCCGTGGGGCGCGTTA
CACCAGGCGA

[000855] The match between the T1 sequence and the C1/C2 sequence is

[000856] Seq. Id. = 117 Position = 47 to 186

[000857] CAGCGTTTTTCTCGCTGTTCTGGACGGCTGAACGCCCTGAATCTCTCCCGG
TATGCAGCCTGCTCGGAGAGTACGATTCGTCGTTGGCTGCACCGAAGTGACGATGGGGCCAT
TCCGTGGGGCGCGTTACACCAGGCGA

[000858] The match between the T2 sequence and the C1/C2 sequence is

[000859] Seq. Id. = 117 Position = 74 to 186

[000860] GCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAGTACGATT
CGTCGTTGGCTGCACCGAAGTGACGATGGGGCCATTCCGTGGGGCGCGTTACACCAGGCGA

[000861] A double stranded DNA loop of length 68.612 kilo-
bases on chromosome 3 (plasmid MP1) is bounded on the left by
a T1 sequence whose identifier is 2692. This T1 control
element has the DNA sequence

[000862] Seq. Id. = 118 Position = 1 to 103

[000863] CTGATGGCCATCCTACAGTACGTTCTCAGCGCGGTCCCGCTGCGCAAGACGC
AGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTTCTCGCTGTTCTTGGAC

[000864] This double stranded DNA loop is bounded on the
right by a T2 control element whose identifier is 2749. This
T2 control element has the DNA sequence

[000865] Seq. Id. = 119 Position = 1 to 103

[000866] AGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTTCTGACCGTGCTGCTCA
GCGTTTTTCTCGCTGTTCTTGGACGGCTGAACGCCCTGAATCTCTCCCGGT

[000867] This long T1/T2 double stranded DNA loop modulates
the expression of the following genes

DRB0059	DRB0060	DRB0061	DRB0062	DRB0064
DRB0065	DRB0066	DRB0067	DRB0068	DRB0069
DRB0070	DRB0072	DRB0073	DRB0074	DRB0076
DRB0077	DRB0079	DRB0080	DRB0081	DRB0083
DRB0085	DRB0086	DRB0087	DRB0088	DRB0089
DRB0090	DRB0092	DRB0093	DRB0094	DRB0096

DRB0097	DRB0098	DRB0102	DRB0103	DRB0104
DRB0105	DRB0106	DRB0107	DRB0111	DRB0112

[000868] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000869] A C1/C2 short loop on chromosome 3 (plasmid MP1) whose identifier is 2693 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene DRB0057 and has the DNA sequence

[000870] Seq. Id. = 120 Position = 1 to 103

[000871] CTGATGGCCATCCTACAGTACGTTCTCAGCGCGGTCCCGCTGCGCAAGACGC
AGCGGAATTCCTGACCGTGCTGCTCAGCGTTTTTCTCGCTGTTCTTGAC

[000872] A C1/C2 short loop on chromosome 3 (plasmid MP1) whose identifier is 2695 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene DRB0057 and has the DNA sequence

[000873] Seq. Id. = 121 Position = 1 to 274

[000874] GCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAGTACGATT
CGTCGTTGGCTGCACCGAAGTGACGATGGGGCCATTCCGTGGGGCGCGTTACACCAGGCGAC
TGTCAGTACAGCAATCGAGAGTGGGCTGATCAGCCCACTGTGCGTTCTGGCCATCGACGCCT
CTTTTCACCGCAAAGCCGGTCAGCACACCGCACACCTCGGCTCGTTCTGGAATGGCTGTGCC
GCGCGGACCGAACGCGGAATCGAGCAATCCTGTTGT

[000875] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[000876] A C1/C2 short loop on chromosome 1 whose identifier is 16 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene DR0009 and has the DNA sequence

[000877] Seq. Id. = 122 Position = 1 to 186

[000878] GCTGTGAAATCACCGCTTCCAATGGGTCTGATGGCCATCCTACAGTACGTTC
TCAGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTT
CTCGCTGTTCTGGACGGCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAG
TACGATTCGT

[000879] The match between the T1 sequence and the C1/C2 sequence is

[000880] Seq. Id. = 122 Position = 28 to 130

[000881] CTGATGGCCATCCTACAGTACGTTCTCAGCGCGGTCCCGCTGCGCAAGACGC
AGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTTCTCGCTGTTCTGGAC

[000882] The match between the T2 sequence and the C1/C2 sequence is

[000883] Seq. Id. = 122 Position = 55 to 157

[000884] AGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTTCTGACCGTGCTGCTCA
GCGTTTTTCTCGCTGTTCTGGACGGCTGAACGCCCTGAATCTCTCCCGGT

[000885] A C1/C2 short loop on chromosome 3 (plasmid MP1) whose identifier is 2768 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene DRB0133 and has the DNA sequence

[000886] Seq. Id. = 123 Position = 1 to 309

[000887] GCTGTGAAATCACCGCTTCCAATGGGTCTGATGGCCATCCTACAGTACGTTC
TCAGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTCCTGACCGTGCTGCTCAGCGTTTTT
CTCGCTGTTCTGGACGGCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGGAGAG
TACGATTCGTCGGACCGAACGCGGAATCGAGCAATCCTGTTGTGCCCTCATTGATGTCCAGC
ACCGGCAGGCCTTGACGGTCGATGTCCGTCAGACCCTGACCGGGTCTGAGGCTCCAACCTCGT
CTGGAACAG

[000888] The match between the T1 sequence and the C1/C2
sequence is

[000889] Seq. Id. = 123 Position = 28 to 130

[000890] CTGATGGCCATCCTACAGTACGTTCTCAGCGCGGTCCCGCTGCGCAAGACGC
AGCGGAATTCCTGACCGTGCTGCTCAGCGTTTTTCTCGCTGTTCTGAC

[000891] The match between the T2 sequence and the C1/C2
sequence is

[000892] Seq. Id. = 123 Position = 55 to 107

[000893] AGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTCCTGACCGTGCTGCTCA
GCGTTTTTCTCGCTGTTCTGACGGCTGAACGCCCTGAATCTCTCCCGGT

[000894] A C1/C2 short loop on chromosome 3 (plasmid MP1)
whose identifier is 2693 controls the expression of the genes
in this T1/T2 long loop. This C1/C2 short loop is expressed
as a RNA single strand that is 3'UTR to the gene DRB0057 and
has the DNA sequence

[000895] Seq. Id. = 124 Position = 1 to 103

[000896] CTGATGGCCATCCTACAGTACGTTCTCAGCGCGGTCCCGCTGCGCAAGACGC
AGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTTCTCGCTGTTCTGAC

[000897] The match between the T1 sequence and the C1/C2
sequence is

[000898] Seq. Id. = 124 Position = 1 to 103

[000899] CTGATGGCCATCCTACAGTACGTTCTCAGCGCGGTCCCGCTGCGCAAGACGC
AGCGGAATTTCTGACCGTGCTGCTCAGCGTTTTTCTCGCTGTTCTGAC

[000900] The match between the T2 sequence and the C1/C2
sequence is

[000901] Seq. Id. = 124 Position = 28 to 103

[000902] AGCGCGGTCCCGCTGCGCAAGACGCAGCGGAATTTCTGACCGTGCTGCTCA
GCGTTTTTCTCGCTGTTCTGAC

[000903] A C1/C2 short loop on chromosome 3 (plasmid MP1)
whose identifier is 2653 controls the expression of the genes
in this T1/T2 long loop. This C1/C2 short loop is expressed
as a RNA single strand that is 3'UTR to the gene DRB0017 and
has the DNA sequence

[000904] Seq. Id. = 125 Position = 1 to 186

[000905] CGGTCCCGCTGCGCAAGACGCAGCGGAATTTCTGACCGTGCTGCTCAGCGT
TTTTCTCGCTGTTCTGACGGCTGAACGCCCTGAATCTCTCCCGGTATGCAGCCTGCTCGG
AGAGTACGATTCGTCGTTGGCTGCACCGAAGTGACGATGGGGCCATTCCGTGGGGCGCGTTA
CACCAGGCGA

[000906] The match between the T1 sequence and the C1/C2
sequence is

[000907] Seq. Id. = 125 Position = 1 to 172

[000908] CGGTCCCGCTGCGCAAGACGCAGCGGAATTCCTGACCGTGCTGCTCAGCGT
TTTTCTCGCTGTTCTGGAC

[000909] The match between the T2 sequence and the C1/C2
sequence is

[000910] Seq. Id. = 125 Position = 1 to 99

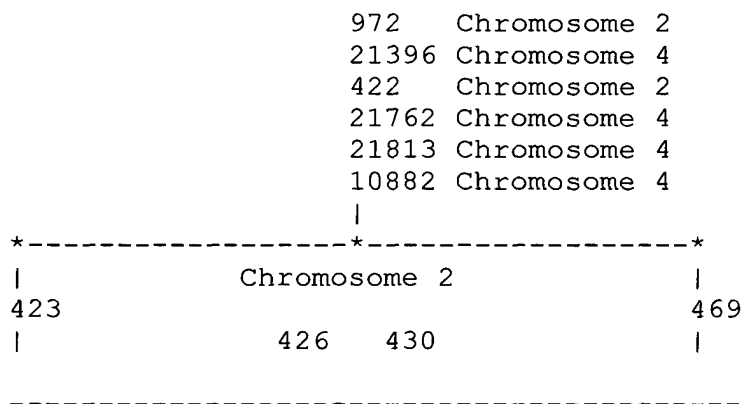
[000911] CGGTCCCGCTGCGCAAGACGCAGCGGAATTCCTGACCGTGCTGCTCAGCGT
TTTTCTCGCTGTTCTGGACGGCTGAACGCCCTGAATCTCTCCCGGT

5. Connectrons occur in plants and higher animals

[000912] Connectron relationships exist in plant and higher animals.

[000913] Example of a plant connectron - *A. thaliana*

[000914] In this example the existence of the T1-T2 (423-469) long loop is controlled by six C1/C2 short loops (972, 21396, 422, 21762, 21813 and 10882). The T1-T2 long loop controls the expression of six genes on chromosome 2 in addition to two C1/C2 (426 and 430) short loops.



[000915] A double stranded DNA loop of length 42.285 kilobases on chromosome 2 is bounded on the left by a T1 sequence whose identifier is 423. This T1 control element has the DNA sequence

[000916] Seq. Id. = 126 Position = 1 to 67

[000917] TATCTCTTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAA
TTAAAAAACGAAATA

[000918] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 469. This T2 control element has the DNA sequence

[000919] Seq. Id. = 127 Position = 1 to 67

[000920] TACTAATTTAATTAATTAATTTAATTTAAAAACGAAATACATTATTAATTT
TCAAAAATAATAACC

[000921] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

At2g02070 At2g02080 At2g02090 At2g02100 At2g02120
At2g02130

[000922] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000923] A C1/C2 short loop on chromosome 2 whose identifier is 426 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene At2g02060 and has the DNA sequence

[000924] Seq. Id. = 128 Position = 1 to 55

[000925] TTCCAAAAATAATAACCAATCAAATCAACATATAAGATTTGATATCTAAAT
TTT

[000926] A C1/C2 short loop on chromosome 2 whose identifier is 430 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene At2g02060 and has the DNA sequence

[000938] TACTAATTTAATTAATTAAATTTAATTAAAAACGAAATACATTATTAATTT

[000939] A C1/C2 short loop on chromosome 4 whose identifier is 21396 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene AT4g15300 and has the DNA sequence

[000940] Seq. Id. = 131 Position = 1 to 122

[000941] TGCCATTAGAAATAAAATTTTAAAGAGTAAATTAATTTATCTCTTTAAGGAT
TAAAAAGTCAAATACTAATTTAATTAATTAATTTAATTAAAAACGAAATACATTATTAAT
TTCCAAAA

[000942] The match between the T1 sequence and the C1/C2 sequence is

[000943] Seq. Id. = 131 Position = 38 to 104

[000944] TATCTCTTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAATTTAA
TTAAAAACGAAATA

[000945] The match between the T2 sequence and the C1/C2 sequence is

[000946] Seq. Id. = 131 Position = 65 to 116

[000947] TACTAATTTAATTAATTAATTTAATTAAAAACGAAATACATTATTAATTT

[000948] A C1/C2 short loop on chromosome 2 whose identifier is 422 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene At2g02060 and has the DNA sequence

[000949] Seq. Id. = 132 Position = 1 to 137

[000950] TAACCTTAATTTTGTAAAGTAATTATATAGGTATGCCATTAGAAATAAAATT
TTAAAGAGTAAATTAATTTATCTCTTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATT
AAATTTAATTAAAAACGAAATA

[000951] The match between the T1 sequence and the C1/C2
sequence is

[000952] Seq. Id. = 132 Position = 71 to 137

[000953] TATCTCTTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAA
TTAAAAACGAAATA

[000954] The match between the T2 sequence and the C1/C2
sequence is

[000955] Seq. Id. = 132 Position = 98 to 137

[000956] TACTAATTTAATTAATTAAATTTAATTAAAAACGAAATA

[000957] A C1/C2 short loop on chromosome 4 whose identifier
is 21762 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene AT4g17510 and has the DNA
sequence

[000958] Seq. Id. = 133 Position = 1 to 65

[000959] TTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAATTAATA
AACGAAATACATT

[000960] The match between the T1 sequence and the C1/C2 sequence is

[000961] Seq. Id. = 133 Position = 1 to 61

[000962] TTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAATTAAAA
AACGAAATA

[000963] The match between the T2 sequence and the C1/C2 sequence is

[000964] Seq. Id. = 133 Position = 22 to 65

[000965] TACTAATTTAATTAATTAAATTTAATTAAAAACGAAATACATT

[000966] A C1/C2 short loop on chromosome 4 whose identifier is 21813 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene AT4g17680 and has the DNA sequence

[000967] Seq. Id. = 134 Position = 1 to 65

[000968] TTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAATTAAAA
AACGAAATACATT

[000969] The match between the T1 sequence and the C1/C2 sequence is

[000970] Seq. Id. = 134 Position = 1 to 61

[000971] TTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAATTAAAA
AACGAAATA

[000972] The match between the T2 sequence and the C1/C2 sequence is

[000973] Seq. Id. = 134 Position = 22 to 65

[000974] TACTAATTTAATTAATTAAATTTAATTAACGAAATACATT

[000975] A C1/C2 short loop on chromosome 2 whose identifier is 10882 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene At2g26540 and has the DNA sequence

[000976] Seq. Id. = 135 Position = 1 to 56

[000977] TATCTCTTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAA
TTAA

[000978] The match between the T1 sequence and the C1/C2 sequence is

[000979] Seq. Id. = 135 Position = 1 to 56

[000980] TATCTCTTTAAGGATTAAAAAGTCAAATACTAATTTAATTAATTAAATTTAA
TTAA

[000981] The match between the T2 sequence and the C1/C2 sequence is

[000982] Seq. Id. = 135 Position = 28 to 56

[000983] TACTAATTTAATTAATTAAATTTAATTAA

[000984] Example of a animal connectron - D. megalomaster

[000985] A double stranded DNA loop of length 88.159 kilobases on chromosome 4 is bounded on the left by a T1 sequence whose identifier is 3340. This T1 control element has the DNA sequence

[000986] Seq. Id. = 136 Position = 1 to 132

[000987] ACCTAAAAGAAGTACCGTTTTTTACTCCTAATTACCAATTCTAACCATCCAT
ATCACTTTTTGACGGACTCCGTGAAAATAATTTTTGGCCAAATTTTCGCATTTTTTGTAAGG
GGTAACATCATAAAAATT

[000988] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3372. This T2 control element has the DNA sequence

[000989] Seq. Id. = 137 Position = 1 to 136

[000990] AAAAAAGTACCGCGTTTTTACTCCTAATTACCAATTCTAACCATCCATATCAC
TTTTTGACGGACTCCGTGAAAATAATTTTTGGCCAAATTTTCGCATTTTTTGTAAGGGGTAA
CATCATCAAAATTTGCGAAAAA

[000991] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

[000992] [Some of the following gene names have not been determined.]

-	-	-	-	-
-	CG11207	-	CG2186	CG2157
-	Ork1	-	-	-
-				

[000993] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[000994] A C1/C2 short loop on chromosome 4 whose identifier is 3362 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene XXX and has the DNA sequence

[000995] Seq. Id. = 138 Position = 1 to 134

[000996] AAAAAAGTACCGCGTTTTACTCCTAATTACCAATTCTAACCATCCATATCAC
TTTTTGACGGACTCCGTTAAAATAATTTTGGACCAAATTTTCGCATTTTTTGTAATCAAAAT
TTGCAAAAAATTGAAAAAAC

[000997] A C1/C2 short loop on chromosome 4 whose identifier is 3364 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene XXX and has the DNA sequence

[000998] Seq. Id. = 139 Position = 1 to 83

[000999] CAAAATTTGAATGCAAATCGATTGGGAATCAAAAAACAACTCAACGAGGTA
TGACATTCCATATTTGGGCCATTATTTCCAA

[0001000] A C1/C2 short loop on chromosome 4 whose identifier is 3366 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene XXX and has the DNA sequence

[0001001] Seq. Id. = 140 Position = 1 to 62

[0001002] TTTTTCACAAAAATTAGGAAAATGATTTTGGGTAAAAAATGAATATTTAA
GTTGGGTTTT

[0001003] A C1/C2 short loop on chromosome 4 whose identifier is 3369 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene XXX and has the DNA sequence

[0001004] Seq. Id. = 141 Position = 1 to 87

[0001005] AAATCGATTGGGAATCAAAAAACAAACCTCAACGAGGTATGACATTCCATAT
CTGGGCCATTATTTCCAATCTTTTGATCAAAATAC

[0001006] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001007] A C1/C2 short loop on chromosome 4 whose identifier is 3373 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene XXX and has the DNA sequence

[0001008] Seq. Id. = 142 Position = 1 to 136

[0001009] AAAAAAGTACCGCGTTTTACTCCTAATTACCAATTCTAACCATCCATATCAC
TTTTTGACGGACTCCGTGAAAATAATTTTTGGCCAAATTTTCGCATTTTTTGTAAGGGGTAA
CATCATCAAAATTTGCGAAAAA

[0001010] The match between the T1 sequence and the C1/C2 sequence is

[0001011] Seq. Id. = 142 Position = 15 to 120

[0001012] TTTTACTCCTAATTACCAATTCTAACCATCCATATCACTTTTTGACGGACTC
CGTGAAAATAATTTTTGGCCAAATTTTCGCATTTTTTGTAAGGGGTAACATCAT

[0001013] The match between the T2 sequence and the C1/C2
sequence is

[0001014] Seq. Id. = 142 Position = 1 to 136

[0001015] AAAAAAGTACCGCGTTTTTACTCCTAATTACCAATTCTAACCATCCATATCAC
TTTTTGACGGACTCCGTGAAAATAATTTTTGGCCAAATTTTCGCATTTTTTGTAAGGGGTAA
CATCATCAAATTTGCGAAAAA

[0001016] Example of an animal connectron - H. sapiens

[0001017] All of the human genome that has been fully
sequenced by both the NIH-lead global sequencing project and
the Celera Genomics, Inc. project. The gene descriptors for
this chromosome do not yet exist. Without the positions and
directions of the genes, it is not possible to select from
among the possible connectrons to determine the real
connectrons.

[0001018] Human chromosome 22 has been processed and there
31,000 possible connectrons.

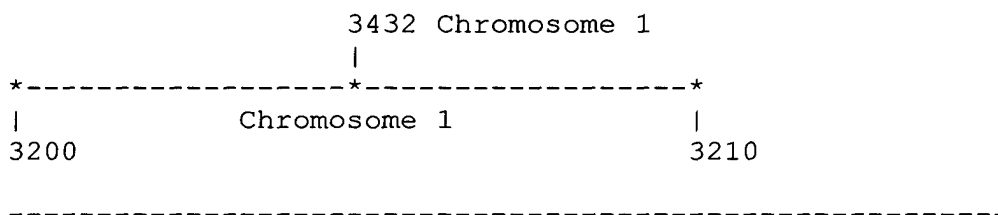
[0001019] The gene descriptors for all the chromosomes of the
human genome should become available within the year.

6. Permanent connectrons exist in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[0001020] C1/C2 short loops are normally expressed as the 3'UTR of some gene. A class of connectron relationships exist that permit one C1/C2 short loop to control the existence of one or more T1-T2 long loops without being subject to any expression controls other than those of the gene to which the C1/C2 is 3'UTR. These connectron relationships are described as "permanent". Permanent connectrons exist in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[0001021] Example of a prokaryote permanent connectron - E. coli

[0001022] In this example the existence of the T1-T2 (3200-3210) long loop is controlled by a C1/C2 short loop (3432). The expression of this C1/C2 short loop is controlled only by the gene *btuB*.



[0001023] A double stranded DNA loop of length 93.339 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 3200. This T1 control element has the DNA sequence

[0001024] Seq. Id. = 143 Position = 1 to 378

[0001025] AAGCGGCACTGCTCTTTAACAATTTATCAGACAATCTGTGTGGGCACTCGAA
GATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGAGTGAACACGTA
ATTCATTACGAAGTTTAATTCTTTGAGCATCAAACCTTTTAAATTGAAGAGTTTGATCATGGC
TCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAACAGCTT
GCTGTTTCGCTGACGAGTGGCGGACGGGTGAGTAATGTCTGGGAACTGCCTGATGGAGGGG
GATAACTACTGGAAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGACCTT
CGGGCCTCTTGCCATC

[0001026] This double stranded DNA loop is bounded on the
right by a T2 control element whose identifier is 3310. This
T2 control element has the DNA sequence

[0001027] Seq. Id. = 144 Position = 1 to 378

[0001028] CAGACAATCTGTGTGGGCACTCGAAGATACGGATTCTTAACGTCGCAAGACG
AAAAATGAATACCAAGTCTCAAGAGTGAACACGTAATTCATTACGAAGTTTAATTCTTTGAG
CGTCAAACCTTTTAAATTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTA
ACACATGCAAGTCGAACGGTAACAGGAAGAAGCTTGCTTCTTTGCTGACGAGTGGCGGACGG
GTGAGTAATGTCTGGGAACTGCCTGATGGAGGGGGATAACTACTGGAAACGGTAGCTAATA
CCGCATAACGTCGCAAGACCAAAGAGGGGGACCTTCGGGCCTCTTGCCATCGGATGTGCCCA
GATGGGATTAGCTAGT

[0001029] This long T1/T2 double stranded DNA loop modulates
the expression of the following genes

rrsC	gltU	rrlC	rrfC	aspT
trpT	yifA	yifE	yifB	ilvL
ilvG_1	ilvM	ilvE	ilvD	ilvA
ilvY	ilvC	ppiC	b3776	rep
gppA	rhlB	trxA	rhoL	rho
rfe	wzzE	wecB	rffH	wecD
wecE	wzxE	yifM_2	wecG	yifK
argX	hisR	leuT	proM	aslB

aslA	hemY	hemX	hemD	cyaA
cyaY	b3808	dapF	uvrD	b3814
corA	yigF	yigG	rarD	yigI
pldA	recQ	yigJ	yigK	pldB
yigL	yigM	metR	metE	ysgA
udp	yigN	ubiE	yigP	b3836
yigU	yigW_1	rfaH	yigC	ubiB
fadA	fadB	pepQ	trkH	hemG

[0001030] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001031] A C1/C2 short loop on chromosome 1 whose identifier is 3432 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene btuB and has the DNA sequence

[0001032] Seq. Id. = 145 Position = 1 to 520

[0001033] AAGCGGCACTGCTCTTTAACAATTTATCAGACAATCTGTGTGGGCACTCGAA
GATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGAGTGAACACGTA
ATTCATTACGAAGTTTAATTCTTTGAGCCAGACAATCTGTGTGGGCACTCGAAGATACGGAT
TCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGAGTGAACACGTAATTCATTAC
GAAGTTTAATTCTTTGAGCGTCAAACCTTTTAAATTGAAGAGTTTGATCATGGCTCAGATTGA
ACGCTGGCGGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAGAAGCTTGCTTCTTTG
CTGACGAGTGGCGGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTAC
TGGAACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGACCTTCGGGCCTCT
TGCCATCGGATGTGCCCAGATGGGATTAGCTAGT

[0001034] The match between the T1 sequence and the C1/C2 sequence is

[0001035] Seq. Id. = 145 Position = 1 to 142

[0001036] AAGCGGCACTGCTCTTTAACAATTTATCAGACAATCTGTGTGGGCACTCGAA
GATACGGATTCTTAACGTCGCAAGACGAAAAATGAATACCAAGTCTCAAGAGTGAACACGTA
ATTCATTACGAAGTTTAATTCTTTGAGC

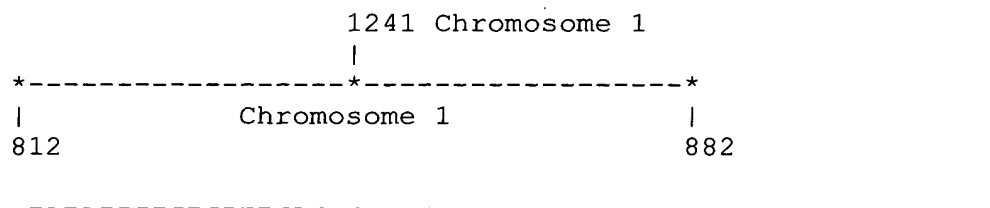
[0001037] The match between the T2 sequence and the C1/C2
sequence is

[0001038] Seq. Id. = 145 Position = 143 to 520

[0001039] CAGACAATCTGTGTGGGCACTCGAAGATACGGATTCTTAACGTCGCAAGACG
AAAAATGAATACCAAGTCTCAAGAGTGAACACGTAATTCATTACGAAGTTTAATTCTTTGAG
CGTCAAACCTTTTAAATTGAAGAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTA
ACACATGCAAGTCGAACGGTAACAGGAAGAAGCTTGCTTCTTTGCTGACGAGTGGCGGACGG
GTGAGTAATGTCTGGGAACTGCCTGATGGAGGGGGATAACTACTGGAAACGGTAGCTAATA
CCGCATAACGTCGCAAGACCAAAGAGGGGGACCTTCGGGCCTCTTGCCATCGGATGTGCCCA
GATGGGATTAGCTAGT

[0001040] Example of an archea permanent connectron - H.
pylori

[0001041] In this example the existence of the T1-T2 (812-
882) long loop is controlled by a C1/C2 short loop (1241).
The expression of this C1/C2 short loop is controlled only by
the gene HP1535.



[0001042] A double stranded DNA loop of length 96.385 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 812. This T1 control element has the DNA sequence

[0001043] Seq. Id. = 146 Position = 1 to 43

[0001044] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCA

[0001045] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 882. This T2 control element has the DNA sequence

[0001046] Seq. Id. = 147 Position = 1 to 43

[0001047] TAGCGGAACTAAAGCATTTCATCCCAAACACTAAAGATATTTGG

[0001048] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

HP0999	HP1000	HP1001	HP1002	HP1003
HP1005	HP1006	HP1008	HP1009	HPtRNA-Pro
HP1010	HP1011	HP1013	HP1015	HP1017
HP1018	HP1020	HP1021	HP1022	HP1023
HP1024	HP1025	HP1027	HP1028	HP1030
HP1031	HP1033	HP1034	HP1038	HP1039
HP1040	HP1041	HP1042	HP1043	HP1044
HP1045	HP1046	HP1051	HP1052	HP1055
HP1056	HP1058	HP1060	HP1065	HPtRNA-Ser
HP1066	HP1067	HP1069	HP1070	HP1074
HP1075	HP1076	HP1077	HP1078	HP1079
HP1080	HP1081	HP1083	HP1084	HP1085
HP1088	HP1091	HP1092	HP1093	HP1094
HP1095	HP1096			

[0001049] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001050] A C1/C2 short loop on chromosome 1 whose identifier is 1241 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene HP1535 and has the DNA sequence

[0001051] Seq. Id. = 148 Position = 1 to 56

[0001052] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCATTCATCCCA
AACA

[0001053] The match between the T1 sequence and the C1/C2 sequence is

[0001054] Seq. Id. = 148 Position = 1 to 43

[0001055] TTTTACTCATAGGGTTTTTATAGTTCCTAGCGGAACTAAAGCA

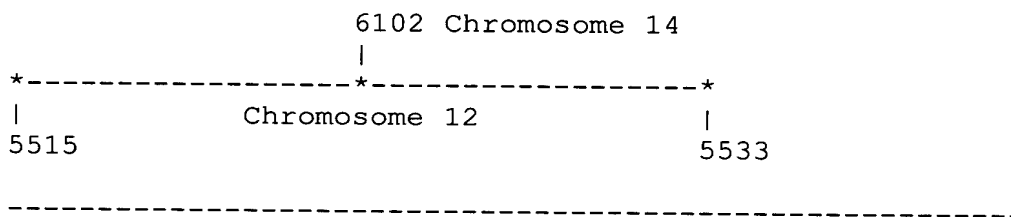
[0001056] The match between the T2 sequence and the C1/C2 sequence is

[0001057] Seq. Id. = 148 Position = 28 to 56

[0001058] TAGCGGAACTAAAGCATTCATCCCAAACA

[0001059] Example of a single-celled permanent connectron - *S. cerevisiae*

[0001060] In this example the existence of the T1-T2 (5515-5533) long loop is controlled by a C1/C2 short loop (6102). The expression of this C1/C2 short loop is controlled only by the gene YNL339C.



[0001061] A double stranded DNA loop of length 6.466 kilobases on chromosome 12 is bounded on the left by a T1 sequence whose identifier is 5515. This T1 control element has the DNA sequence

[0001062] Seq. Id. = 149 Position = 1 to 225

[0001063] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGTTG

[0001064] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 5533. This T2 control element has the DNA sequence

[0001065] Seq. Id. = 150 Position = 1 to 225

[0001066] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTTAGTGTTTGTGTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGTTGAACATCCGGGTAAGAGACAACAGGGCT

[0001067] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

YLR467W

[0001068] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001069] A C1/C2 short loop on chromosome 14 whose identifier is 6102 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YNL339C and has the DNA sequence

[0001070] Seq. Id. = 151 Position = 1 to 252

[0001071] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAA
GAGACAACAGGGCT

[0001072] The match between the T1 sequence and the C1/C2 sequence is

[0001073] Seq. Id. = 151 Position = 1 to 225

[0001074] AGGAAATTGTTGTTACGAAAGTCAGTGATTATGTATTGTGTAGTATAGTATA
TTGTAAGAAATTTTTTTTCTAGGGAATATGCGTTTTGATGTAGTAGTATTTCACTGTTTTG
ATTTAGTGTTTGTGTCACGGCAGTAGCGAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACA
ATCTATAAAAAGTAAACATAAAATAAAGGTAGTAAGTAGCTTTTGGTTG

[0001075] The match between the T2 sequence and the C1/C2 sequence is

[0001076] Seq. Id. = 151 Position = 28 to 252

[0001077] ATTATGTATTGTGTAGTATAGTATATTGTAAGAAATTTTTTTTCTAGGGAA
TATGCGTTTTGATGTAGTAGTATTTCACTGTTTTGATTAGTGTTTGTGTCACGGCAGTAGC
GAGAGACAAGTGGGAAAGAGTAGGATAAAAAGACAATCTATAAAAAGTAAACATAAAATAAA
GGTAGTAAGTAGCTTTTGGTTGAACATCCGGGTAAGAGACAACAGGGCT

[0001078] Example of a multi-celled permanent connectron - C.
elegans

[0001079] In this example the existence of the T1-T2 (5515-
5533) long loop is controlled by a C1/C2 short loop (6102).
The expression of this C1/C2 short loop is controlled only by
the gene YNL339C.

24442 Chromosome 5
|

| Chromosome 1 |
569 596

[0001080] A double stranded DNA loop of length 30.606 kilo-
bases on chromosome 1 is bounded on the left by a T1 sequence
whose identifier is 569. This T1 control element has the DNA
sequence

[0001081] Seq. Id. = 152 Position = 1 to 239

[0001082] AAATCGAGCCCGTAAATCGACACAAGCGCTACAGTAGTC

[0001083] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 596. This T2 control element has the DNA sequence

[0001084] Seq. Id. = 153 Position = 1 to 42

[0001085] AGTGCTACAGTAGTCATTTAAAGAATTACTGTAGTTTTTCGCT

[0001086] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001087] A C1/C2 short loop on chromosome 5 whose identifier is 24442 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene F20D6.4 and has the DNA sequence

[0001088] Seq. Id. = 154 Position = 1 to 58

[0001089] GAGCCCGTAAATCGACACAAGCGCTACAGTAGTCATTTAAAGAATTACTGTAGTTTTTC

[0001090] The match between the T1 sequence and the C1/C2 sequence is

[0001091] Seq. Id. = 154 Position = 1 to 34

[0001092] GAGCCCGTAAATCGACACAAGCGCTACAGTAGTC

[0001093] The match between the T2 sequence and the C1/C2 sequence is

[0001094] Seq. Id. = 154 Position = 23 to 58

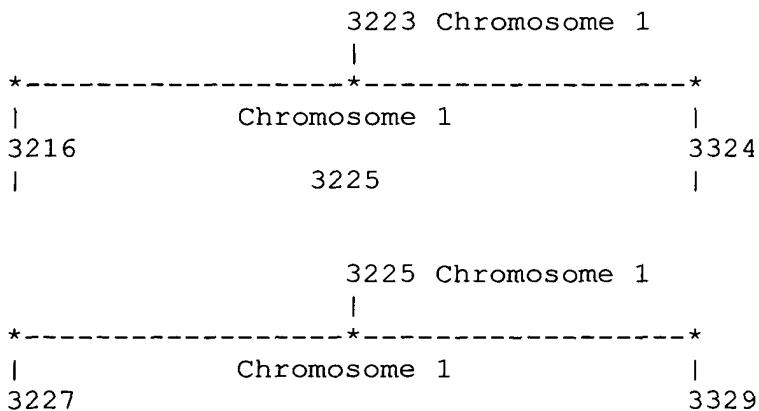
[0001095] GCTACAGTAGTCATTTAAAGAATTACTGTAGTTTTC

7. Transient connectrons exist in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[0001096] A class of connectron relationships exist that permit one C1/C2 short loop to control the existence of one or more T1-T2 long loops such that this C1/C2 short loop is itself subject to expression control by another T1-T2 long loop which surrounds it. These connectron relationships are described as "transient". Transient connectrons exist in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[0001097] Example of a prokaryote transient connectron - E. coli

[0001098] In this example the existence of the T1-T2 (3227-3329) long loop is controlled by the C1/C2 (3225) short loop. The expression of this C1/C2 short loop is controlled by the existence of the T1-T2 (3216-3224) long loop. The existence of this T1-T2 long loop is itself determined by the expression of the C1/C2 (3223) short loop. The C1/C2 (3225) short loop is the transient connectron.



[0001099] A double stranded DNA loop of length 93.464 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 3216. This T1 control element has the DNA sequence

[0001100] Seq. Id. = 155 Position = 1 to 337

[0001101] AGCGCAAGCGAAGCTCTTGATCGAAGCCCCGGTAAACGGCGGCCGTAACCTAT
AACGGTCCTAAGGTAGCGAAATTCCTTGTCGGGTAAGTCCGACCTGCACGAATGGCGTAAT
GATGGCCAGGCTGTCTCCACCCGAGACTCAGTGAAATTGAACTCGCTGTGAAGATGCAGTGT
ACCCGCGGCAAGACGGAAAGACCCCGTGAACCTTTACTATAGCTTGACACTGAACATTGAGC
CTTGATGTGTAGGATAGGTGGGAGGCTTTGAAGTGTGGACGCCAGTCTGCATGGAGCCGACC
TTGAAATACCACCCTTTAATGTTTGATGTTCTAACGT

[0001102] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3324. This T2 control element has the DNA sequence

[0001103] Seq. Id. = 156 Position = 1 to 337

[0001104] CCCGGTAAACGGCGGCCGTAACCTATAACGGTCCTAAGGTAGCGAAATTCCTT
GTCGGGTAAGTTCCGACCTGCACGAATGGCGTAATGATGGCCAGGCTGTCTCCACCCGAGAC
TCAGTGAAATTGAACTCGCTGTGAAGATGCAGTGTACCCGCGGCAAGACGGAAAGACCCCGT
GAACCTTTACTATAGCTTGACACTGAACATTGAGCCTTGATGTGTAGGATAGGTGGGAGGCT
TTGAAGTGTGGACGCCAGTCTGCATGGAGCCGACCTTGAAATACCACCCTTTAATGTTTGAT
GTTCTAACGTTGACCCGTAATCCGGGTTGCGGACAGT

[0001105] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

rrfC	aspT	trpT	yifA	yifE
yifB	ilvL	ilvG_1	ilvM	ilvE

ilvD	ilvA	ilvY	ilvC	ppiC
b3776	rep	gppA	rhlB	trxA
rhoL	rho	rfe	wzzE	wecB
rffH	wecD	wecE	wzxE	yifM_2
wecG	yifK	argX	hisR	leuT
proM	aslB	aslA	hemY	hemX
hemD	cyaA	cyaY	b3808	dapF
uvrD	b3814	corA	yigF	yigG
rarD	yigI	pldA	recQ	yigJ
yigK	pldB	yigL	yigM	metR
metE	ysgA	udp	yigN	ubiE
yigP	b3836	yigU	yigW_1	rfaH
yigC	ubiB	fadA	fadB	pepQ
trkH	hemG	rrsA	ileT	rrlA

[0001106] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001107] A C1/C2 short loop on chromosome 1 whose identifier is 3225 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrlC and has the DNA sequence

[0001108] Seq. Id. = 157 Position = 1 to 137

[0001109] AAACAGAATTTGCCTGGCGGCCGTAGCGCGGTGGTCCCACCTGACCCCATGC
CGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTA
GGGAACTGCCAGGCATCAAATTA

[0001110] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001111] A C1/C2 short loop on chromosome 1 whose identifier is 3323 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrlA and has the DNA sequence

[0001112] Seq. Id. = 158 Position = 1 to 362

[0001113] GCGAAGCTCTTGATCGAAGCCCCGGTAAACGGCGGCCGTAAC TATAACGGTC
CTAAGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCTGCACGAATGGCGTAATGATGGCC
AGGCTGTCTCCACCCGAGACTCAGTGAAATTGAACTCGCTGTGAAGATGCAGTGTACCCGCG
GCAAGACGGAAAGACCCCGTGAACCTTTACTATAGCTTGACACTGAACATTGAGCCTTGATG
TG TAGGATAGGTGGGAGGCTTTGAAGTGTGGACGCCAGTCTGCATGGAGCCGACCTTGAAAT
ACCACCCTTTAATGTTTGATGTTCTAACGTAACGTTGACCCGTAATCCGGGTGCGGACAGT

[0001114] The match between the T1 sequence and the C1/C2 sequence is

[0001115] Seq. Id. = 158 Position = 1 to 330

[0001116] GCGAAGCTCTTGATCGAAGCCCCGGTAAACGGCGGCCGTAAC TATAACGGTC
CTAAGGTAGCGAAATTCCTTGTCGGGTAAGTTCCGACCTGCACGAATGGCGTAATGATGGCC
AGGCTGTCTCCACCCGAGACTCAGTGAAATTGAACTCGCTGTGAAGATGCAGTGTACCCGCG
GCAAGACGGAAAGACCCCGTGAACCTTTACTATAGCTTGACACTGAACATTGAGCCTTGATG
TG TAGGATAGGTGGGAGGCTTTGAAGTGTGGACGCCAGTCTGCATGGAGCCGACCTTGAAAT
ACCACCCTTTAATGTTTGATGTTCTAACGT

[0001117] The match between the T2 sequence and the C1/C2 sequence is

[0001118] Seq. Id. = 158 Position = 21 to 362

[0001119] CCCGGTAAACGGCGGCCGTAAC TATAACGGTCCTAAGGTAGCGAAATTCCTT
GTCGGGTAAGTTCCGACCTGCACGAATGGCGTAATGATGGCCAGGCTGTCTCCACCCGAGAC
TCAGTGAAATTGAACTCGCTGTGAAGATGCAGTGTACCCGCGGCAAGACGGAAAGACCCCGT

GAACCTTTACTATAGCTTGACACTGAACATTGAGCCTTGATGTGTAGGATAGGTGGGAGGCT
 TTGAAGTGTGGACGCCAGTCTGCATGGAGCCGACCTTGAAATACCACCCTTTAATGTTTGAT
 GTTCTAACGTTGACCCGTAATCCGGGTTGCGGACAGT

[0001120] A double stranded DNA loop of length 93.749 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 3227. This T1 control element has the DNA sequence

[0001121] Seq. Id. = 159 Position = 1 to 52

[0001122] AGCGCCGATGGTAGTGTGGGGTCTCCCATGCGAGAGTAGGGAACTGCCAGG

[0001123] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3329. This T2 control element has the DNA sequence

[0001124] Seq. Id. = 160 Position = 1 to 52

[0001125] CATGCGAGAGTAGGGAACTGCCAGGCATCAAATAAAACGAAAGGCTCAGTCG

[0001126] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

aspT	trpT	yifA	yifE	yifB
ilvL	ilvG_1	ilvM	ilvE	ilvD
ilvA	ilvY	ilvC	ppiC	b3776
rep	gppA	rhlB	trxA	rhoL
rho	rfe	wzzE	wecB	rffH
wecD	wecE	wzxE	yifM_2	wecG
yifK	argX	hisR	leuT	proM
aslB	aslA	hemY	hemX	hemD

cyaA	cyaY	b3808	dapF	uvrD
b3814	corA	yigF	yigG	rarD
yigI	pIdA	recQ	yigJ	yigK
pIdB	yigL	yigM	metR	metE
ysgA	udp	yigN	ubiE	yigP
b3836	yigU	yigW_1	rfaH	yigC
ubiB	fadA	fadB	pepQ	trkH
hemG	rrsA	ileT	rrlA	rrfA

[0001127] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001128] A C1/C2 short loop on chromosome 1 whose identifier is 3225 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene rrlC and has the DNA sequence

[0001129] Seq. Id. = 161 Position = 1 to 137

[0001130] AACAGAATTTGCCTGGCGGCCGTAGCGCGGTGGTCCCACCTGACCCCATGC
CGAACTCAGAAGTGAAACGCCGTAGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTA
GGGAACTGCCAGGCATCAAATTA

[0001131] The match between the T1 sequence and the C1/C2 sequence is

[0001132] Seq. Id. = 161 Position = 76 to 127

[0001133] AGCGCCGATGGTAGTGTGGGGTCTCCCCATGCGAGAGTAGGGAACTGCCAGG

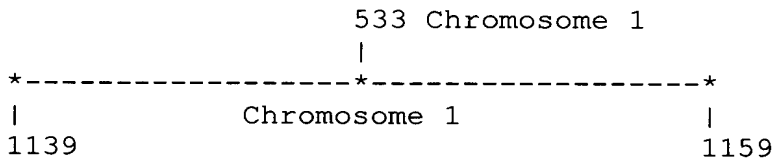
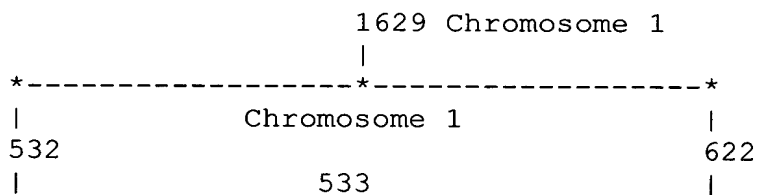
[0001134] The match between the T2 sequence and the C1/C2 sequence is

[0001135] Seq. Id. = 161 Position = 103 to 135

[0001136] CATGCGAGAGTAGGGAACTGCCAGGCATCAAAT

[0001137] Example of an archea transient connectron - M. jannaschii

[0001138] In this example the existence of the T1-T2 (1139-1159) long loop is controlled by the C1/C2 (533) short loop. The expression of this C1/C2 short loop is controlled by the existence of the T1-T2 (532-622) long loop. The existence of this T1-T2 long loop is itself determined by the expression of the C1/C2 (1629) short loop. The C1/C2 (533) short loop is the transient connectron.



[0001139] A double stranded DNA loop of length 78.672 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 532. This T1 control element has the DNA sequence

[0001140] Seq. Id. = 162 Position = 1 to 33

[0001141] ATATGTTTGAAATTTGAAAATAAGAGTATTTAG

[0001142] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 622. This T2 control element has the DNA sequence

[0001143] Seq. Id. = 163 Position = 1 to 47

[0001144] TTGAAAATAAGAGCATTTAGAAAGTTATTAATTAGTTCAAAGGATTTT

[0001145] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

MJ0486	MJ0487	MJ0488	MJ0489	MJ0490
MJ0492	MJ0493	MJ0494	MJ0495	MJ0496
MJ0497	MJ0499	MJ0500	MJ0501	MJ0502
MJ0503	MJ0504	MJ0506	MJ0507	MJ0508
MJ0509	MJ0510	MJ0511	MJ0512	MJ0513
MJ0514	MJ0514	MJ0517	MJ0519	MJ0520
MJ0521	MJ0522	MJ0523	MJ0525	MJ0526
MJ0526	MJ0529	MJ0530	MJ0531	MJ0532
MJ0534	MJ0535	MJ0536	MJ0538	MJ0539
MJ0540	MJ0541	MJ0542	MJ0543	MJ0544
MJ0545	MJ0547	MJ0548	MJ0549	MJ0550
MJ0552	MJ0553	MJ0554	MJ0555	MJ0556
MJ0557	MJ0558	MJ0559	MJ0560	MJ0561
MJ0562	MJ0563	MJ0564	MJ0565	MJ0566
MJ0568	MJ0569	MJ0570		

[0001146] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001147] A C1/C2 short loop on chromosome 1 whose identifier is 533 controls the expression of the genes of one or more

other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ0485 and has the DNA sequence

[0001148] Seq. Id. = 164 Position = 1 to 64

[0001149] ATTTTATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGA
GTTTATTGAATT

[0001150] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001151] A C1/C2 short loop on chromosome 1 whose identifier is 1629 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1597 and has the DNA sequence

[0001152] Seq. Id. = 165 Position = 1 to 139

[0001153] ATATGTTTGAAATTTGAAAATAAGAGTATTTAGAAGTTATTAATTAGTTCAA
AGGATTTTATTTAATTTCTAAGGGTTTGCTGGTTTGATTATTTAGAATATTTGAGTTTATT
GAATTATTCAGATTTTAAAAATTA

[0001154] The match between the T1 sequence and the C1/C2 sequence is

[0001155] Seq. Id. = 165 Position = 1 to 33

[0001156] ATATGTTTGAAATTTGAAAATAAGAGTATTTAG

[0001157] The match between the T2 sequence and the C1/C2 sequence is

[0001158] Seq. Id. = 165 Position = 33 to 60

[0001159] ATTTAGAAGTTATTAATTAGTTCAAAGGATTTT

[0001160] A double stranded DNA loop of length 14.509 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 1139. This T1 control element has the DNA sequence

[0001161] Seq. Id. = 166 Position = 1 to 78

[0001162] ATTTATTAATTAGTTCAAAGGATTTTATTTAATTTCTAAGGGTTAGCTGGT
TTGATTGTTTAAAATATTTGAGTTTA

[0001163] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1159. This T2 control element has the DNA sequence

[0001164] Seq. Id. = 167 Position = 1 to 78

[0001165] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTAT
TGAATTATTCAGATTTTAAAAATTA

[0001166] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

MJ1096	MJ1097	tRNA-Arg-3	MJ1098	MJ1099
MJ1100	MJ1101	MJ1102	MJ1103	MJ1104
MJ1105	MJ1106	MJ1107	MJ1108	

[0001167] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001168] A C1/C2 short loop on chromosome 1 whose identifier is 533 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ0485 and has the DNA sequence

[0001169] Seq. Id. = 168 Position = 1 to 64

[0001170] ATTTTATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTATTGAATT

[0001171] The match between the T1 sequence and the C1/C2 sequence is

[0001172] Seq. Id. = 168 Position = 1 to 37

[0001173] ATTTTATTTAATTTCTAAGGGTTAGCTGGTTTGATT

[0001174] The match between the T2 sequence and the C1/C2 sequence is

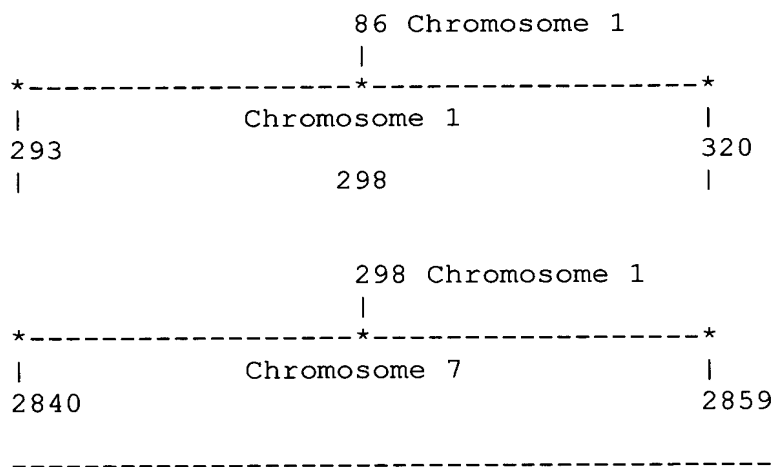
[0001175] Seq. Id. = 168 Position = 7 to 64

[0001176] ATTTAATTTCTAAGGGTTAGCTGGTTTGATTATTTAGAATATTTGAGTTTATTGAATT

[0001177] Example of a single-celled transient connectron - *S. cerevisiae*

[0001178] In this example the existence of the T1-T2 (2840-2859) long loop is controlled by the C1/C2 (298) short loop.

The expression of this C1/C2 short loop is controlled by the existence of the T1-T2 (293-320) long loop. The existence of this T1-T2 long loop is itself determined by the expression of the C1/C2 (86) short loop. The C1/C2 (298) short loop is the transient connectron.



[0001179] A double stranded DNA loop of length 38.470 kilobases on chromosome 2 is bounded on the left by a T1 sequence whose identifier is 293. This T1 control element has the DNA sequence

[0001180] Seq. Id. = 169 Position = 1 to 258

[0001181] GAATTGTTGGAATAAAAAATCCACTATCGTCTATCAACTAATAGTTATATTAT
 CAATATATTATCATATACGGTGTTAAGATGATGACATAAGTTATGAGAAGCTGTCATCGAAG
 TTAGAGGAAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAATGAAACATATAAAACGGA
 ATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATAT
 CCTTGAGGAGAACTTCTAGT

[0001182] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 320. This T2 control element has the DNA sequence

[0001183] Seq. Id. = 170 Position = 1 to 70

[0001184] AATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTCG
AGGAGAACTTCTAGTATATTCTGTA

[0001185] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

YBL005W-B	TS(AGA)B	YBL004W	YBL003C	YBL002W
YBL001C	YBR001C	YBR002C	YBR003W	YBR004C
YBR005W	YBR006W	YBR007C	YBR008C	YBR009C
YBR010W	YBR011C	YBR012C		

[0001186] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001187] A C1/C2 short loop on chromosome 2 whose identifier is 298 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YBL005W-B and has the DNA sequence

[0001188] Seq. Id. = 171 Position = 1 to 342

[0001189] ATCTATTACATTATGGGTGGTATGTTGGAATAAAAATCCACTATCGTCTATC
AACTAATAGTTATATTATCAATATATTATCATATACGGTGTTAAGATGATGACATAAGTTAT
GAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAA
TGAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCA
TTTTGAGGATTCCTATATCCTTGAGGAGAACTTCTAGTATATTCTGTATACCTAATATTATA
GCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTC

[0001190] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001191] A C1/C2 short loop on chromosome 1 whose identifier is 86 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YAR009C and has the DNA sequence

[0001192] Seq. Id. = 172 Position = 1 to 362

[0001193] ATCTATTACATTATGGGTGGTATGTTGGAATAGAAATCAACTATCATCTACT
AACTAGTATTTACATTACTAGTATATTATCATATACGGTGTTAGAAGATGACGCAAATGATG
AGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAA
TGAATATAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGA
TTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAATA
TTATAGCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTCACCCATTTCTCAGAA

[0001194] The match between the T1 sequence and the C1/C2 sequence is

[0001195] Seq. Id. = 172 Position = 184 to 264

[0001196] AAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATAT
AGATTCCATTTTGAGGATTCCTATATCCT

[0001197] The match between the T2 sequence and the C1/C2 sequence is

[0001198] Seq. Id. = 172 Position = 215 to 291

[0001199] AATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTCG
AGGAGAACTTCTAGTATATTCTGTA

[0001200] A double stranded DNA loop of length 5.302 kilobases on chromosome 7 is bounded on the left by a T1 sequence whose identifier is 2840. This T1 control element has the DNA sequence

[0001201] Seq. Id. = 173 Position = 1 to 313

[0001202] TCTGTTGGAATAAAAATCCACTATCGTCTATCAACTAATAGTTATATTATCA
ATATATTATCATATACGGTGTTAAGATGATGACATAAGTTATGAGAAGCTGTCATCGAAGTT
AGAGGAAGCTGAAACGCAAGGATTGATAATGTAATAGGATCAATGAATATAAACATATAAAA
CGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCT
ATATCCTCGAGGAGAACTTCTAGTATATTCTGTATACCTAAATTATAGCCTTTATCAACAAT
GGAATCCCAACAA

[0001203] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 2859. This T2 control element has the DNA sequence

[0001204] Seq. Id. = 174 Position = 1 to 314

[0001205] CTATCAACTAATAGTTATATTATCAATATATTATCATATACGGTGTTAAGAT
GATGACATAAGTTATGAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAAACGCAAGGATTGAT
AATGTAATAGGATCAATGAATATAAACATATAAAACGGAATGAGGAATAATCGTAATATTAG
TATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTCGAGGAGAACTTCTAGTATA
TTCTGTATACCTAATATTATAGCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATT
CACATATTTCTCAT

[0001206] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001207] A C1/C2 short loop on chromosome 2 whose identifier is 298 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YBL005W-B and has the DNA sequence

[0001208] Seq. Id. = 175 Position = 1 to 342

[0001209] ATCTATTACATTATGGGTGGTATGTTGGAATAAAAAATCCACTATCGTCTATC
AACTAATAGTTATATTATCAATATATTATCATATACGGTGTTAAGATGATGACATAAGTTAT
GAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAA
TGAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCA
TTTTGAGGATTCCTATATCCTTGAGGAGAACTTCTAGTATATTCTGTATACCTAATATTATA
GCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTC

[0001210] The match between the T1 sequence and the C1/C2 sequence is

[0001211] Seq. Id. = 175 Position = 23 to 147

[0001212] TGTTGGAATAAAAAATCCACTATCGTCTATCAACTAATAGTTATATTATCAAT
ATATTATCATATACGGTGTTAAGATGATGACATAAGTTATGAGAAGCTGTCATCGAAGTTAG
AGGAAGCTGAA

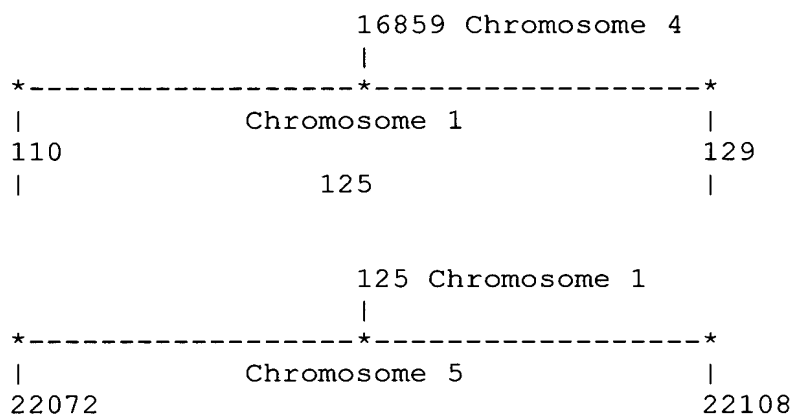
[0001213] The match between the T2 sequence and the C1/C2 sequence is

[0001214] Seq. Id. = 175 Position = 48 to 146

[0001215] CTATCAACTAATAGTTATATTATCAATATATTATCATATACGGTGTTAAGAT
GATGACATAAGTTATGAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAA

[0001216] Example of a multi-celled transient connectron - C. elegans

[0001217] In this example the existence of the T1-T2 (22072-22108) long loop is controlled by the C1/C2 (125) short loop. The expression of this C1/C2 short loop is controlled by the existence of the T1-T2 (110-129) long loop. The existence of this T1-T2 long loop is itself determined by the expression of the C1/C2 (16859) short loop. The C1/C2 (125) short loop is the transient connectron.



[0001218] A double stranded DNA loop of length 18.855 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 110. This T1 control element has the DNA sequence

[0001219] Seq. Id. = 176 Position = 1 to 33

[0001220] AGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGGC

[0001221] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 129. This T2 control element has the DNA sequence

[0001222] Seq. Id. = 177 Position = 1 to 2123

[0001223] TTCTCCCGCATTTTTTGTAGATCTACGTAGATCAAACCGAAATGAGGCACTT
TCTGAATCCACGAGCTAGGCTTAAGCTTAGGCTTAAGCTTAGGCCTTTTCTCAGGCTTAGGC
TTAGGCTTA

[0001224] This long T1/T2 double stranded DNA loop modulates
the expression of the following genes

ZC123.3 ZC123.2

[0001225] This long T1/T2 double stranded DNA loop modulates
the expression of the following C1/C2 short loops

[0001226] A C1/C2 short loop on chromosome 1 whose identifier
is 125 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop is expressed as
a RNA single strand that is 3'UTR to the gene ZC123.3 and has
the DNA sequence

[0001227] Seq. Id. = 178 Position = 1 to 89

[0001228] ACGCGCCGTAAATCTACCCCAGATATGGCCGAGCCAAAATGGCCTAGTTCGG
CAAACCTCTTTCATTTCAATTTATGAGGGAAGCCAGAA

[0001229] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[0001230] A C1/C2 short loop on chromosome 4 whose identifier
is 16859 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene F58E2.7 and has the DNA
sequence

[0001231] Seq. Id. = 179 Position = 1 to 166

[0001232] CTTAGGCTTAAGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGGCTTAAGCTTA
GGCTTAAGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGG
CTTAAGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGACTTA

[0001233] The match between the T1 sequence and the C1/C2
sequence is

[0001234] Seq. Id. = 179 Position = 11 to 43

[0001235] AGCTTAGGCTTAAGCTTAGGCTTAAGCTTAGGC

[0001236] The match between the T2 sequence and the C1/C2
sequence is

[0001237] Seq. Id. = 179 Position = 3 to 33

[0001238] TAGGCTTAAGCTTAGGCTTAAGCTTAGGC

[0001239] A double stranded DNA loop of length 51.031 kilo-
bases on chromosome 5 is bounded on the left by a T1 sequence
whose identifier is 22072. This T1 control element has the
DNA sequence

[0001240] Seq. Id. = 180 Position = 1 to 57

[0001241] CGCAACGCGCCGTAAATCTACCCCAGATATGGCCGAGCCAAAATGACCTAGT
TCGGC

[0001242] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 22108. This T2 control element has the DNA sequence

[0001243] Seq. Id. = 181 Position = 1 to 170

[0001244] TGACAATCGCCTGCCGGACAACGCGTGGAAGTGTGCTGTACTCCACACGG
ACAAATACATTTAGTTTTACAATAAAATCGAACCGCGACGCGACACGCAACGCGACGTAAA
TCTACCCCAGATATGGCCGAGCCAAAATGGCCTAGTTCGGCAAACCTCTTCTATTTTC

[0001245] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

F36H9.3 F36H9.4 F36H9.5 F36H9.2 F36H9.1
F36H9.6

[0001246] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001247] A C1/C2 short loop on chromosome 1 whose identifier is 125 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZC123.3 and has the DNA sequence

[0001248] Seq. Id. = 182 Position = 1 to 89

[0001249] ACGCGCCGTAAATCTACCCCAGATATGGCCGAGCCAAAATGGCCTAGTTCGG
CAAACCTCTTTCATTTCAATTTATGAGGGAAGCCAGAA

[0001250] The match between the T1 sequence and the C1/C2 sequence is

[0001251] Seq. Id. = 182 Position = 1 to 41

[0001252] ACGCGCCGTAAATCTACCCCAGATATGGCCGAGCCAAAATG

[0001253] The match between the T2 sequence and the C1/C2 sequence is

[0001254] Seq. Id. = 182 Position = 7 to 61

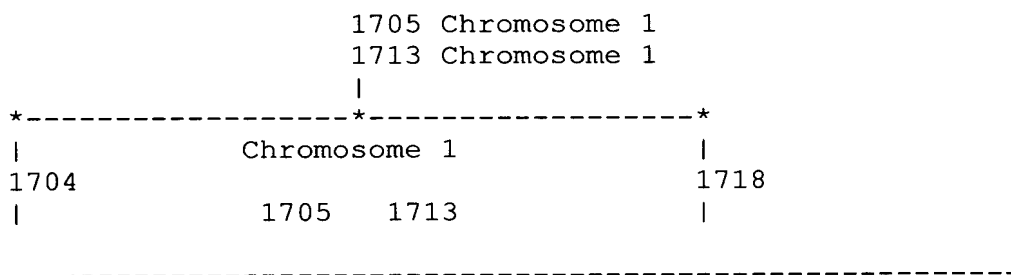
[0001255] CGTAAATCTACCCCAGATATGGCCGAGCCAAAATGGCCTAGTTCGGCAAAC
CTT

8. Self-limiting connectrons occur in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes

[0001256] A class of connectron relationships exist that permit one C1/C2 short loop to control the existence of the T1-T2 long loop that surrounds it. These connectron relationships are described as "self-limiting". Self-limiting connectrons exist in prokaryotes, archaea, single-celled eukaryotes and multi-celled eukaryotes.

[0001257] Example of a prokaryotic self-limiting connectrons - E. coli

[0001258] In this example the existence of the T1-T2 (1704-1718) long loop is controlled by two C1/C2 (1705 and 1713) short loops. The expression of these C1/C2 short loops is controlled by the existence of the T1-T2 (1704-1718) long loop. The existence of this T1-T2 long loop is itself determined by the expression of the two C1/C2 (1705 and 1713) short loops. The C1/C2 (1705 and 1713) short loops are the self-limiting connectrons.



[0001259] A double stranded DNA loop of length 15.259 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 1704. This T1 control element has the DNA sequence

[0001260] Seq. Id. = 183 Position = 1 to 71

[0001261] CGCCCCGTTACACGATTCCTCTGTAGTTCAGTCGGTAGAACGGCGGACTGT
TAATCCGTATGTCACTGGT

[0001262] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1718. This T2 control element has the DNA sequence

[0001263] Seq. Id. = 184 Position = 1 to 71

[0001264] TTCAGTCGGTAGAACGGCGGACTGTTAATCCGTATGTCACTGGTTCGAGTCC
AGTCAGAGGAGCCAAATTC

[0001265] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

asnT	b1978	b1979	b1980	shiA
amn	b1983	asnW	yeeO	asnU

[0001266] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001267] A C1/C2 short loop on chromosome 1 whose identifier is 1705 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene and has the DNA sequence

[0001268] Seq. Id. = 185 Position = 1 to 98

[0001269] CGCCCCGTTACACGATTCCTCTGTAGTTCAGTCGGTAGAACGGCGGACTGT
TAATCCGTATGTCACTGGTTCGAGTCCAGTCAGAGGAGCCAAATTC

[0001270] A C1/C2 short loop on chromosome 1 whose identifier is 1713 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene asnW and has the DNA sequence

[0001271] Seq. Id. = 186 Position = 1 to 86

[0001272] CACGATTCCTCTGTAGTTCAGTCGGTAGAACGGCGGACTGTTAATCCGTATG
TCACTGGTTCGAGTCCAGTCAGAGGAGCCAAATT

[0001273] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001274] A C1/C2 short loop on chromosome 1 whose identifier is 1705 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene and has the DNA sequence

[0001275] Seq. Id. = 187 Position = 1 to 98

[0001276] CGCCCCGTTACACGATTCCTCTGTAGTTCAGTCGGTAGAACGGCGGACTGT
TAATCCGTATGTCACTGGTTCGAGTCCAGTCAGAGGAGCCAAATTC

[0001277] The match between the T1 sequence and the C1/C2 sequence is

[0001278] Seq. Id. = 187 Position = 1 to 71

[0001279] CGCCCCGTTACACGATTCCTCTGTAGTTCAGTCGGTAGAACGGCGGACTGT
TAATCCGTATGTCACTGGT

[0001280] The match between the T2 sequence and the C1/C2 sequence is

[0001281] Seq. Id. = 187 Position = 28 to 98

[0001282] TTCAGTCGGTAGAACGGCGGACTGTTAATCCGTATGTCAGTGGTTCGAGTCC
AGTCAGAGGAGCCAAATTC

[0001283] A C1/C2 short loop on chromosome 1 whose identifier is 1713 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene asnW and has the DNA sequence

[0001284] Seq. Id. = 188 Position = 1 to 86

[0001285] CACGATTCCTCTGTAGTTCAGTCGGTAGAACGGCGGACTGTTAATCCGTATG
TCACTGGTTCGAGTCCAGTCAGAGGAGCCAAATT

[0001286] The match between the T1 sequence and the C1/C2 sequence is

[0001287] Seq. Id. = 188 Position = 1 to 60

[0001288] CACGATTCCTCTGTAGTTCAGTCGGTAGAACGGCGGACTGTTAATCCGTATG
TCACTGGT

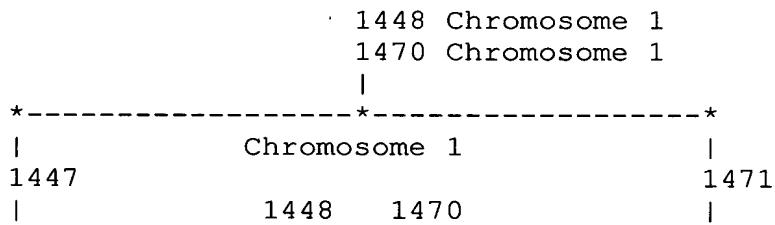
[0001289] The match between the T2 sequence and the C1/C2 sequence is

[0001290] Seq. Id. = 188 Position = 17 to 86

[0001291] TTCAGTCGGTAGAACGGCGGACTGTTAATCCGTATGTCAGTGGTTCGAGTCC
AGTCAGAGGAGCCAAATT

[0001292] Example of a archea self-limiting connectrons - M. jannaschii

[0001293] In this example the existence of the T1-T2 (1447-1471) long loop is controlled by two C1/C2 (1448 and 1470) short loops. The expression of these C1/C2 short loops is controlled by the existence of the T1-T2 (1447-1471) long loop. The existence of this T1-T2 long loop is itself determined by the expression of the two C1/C2 (1705 and 1713) short loops. The C1/C2 (1448 and 1470) short loops are the self-limiting connectrons.



[0001294] A double stranded DNA loop of length 22.675 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 1447. This T1 control element has the DNA sequence

[0001295] Seq. Id. = 189 Position = 1 to 95

[0001296] TTATAGAACATTATGAAGCTTTTTACTCAACTAACAACCGTATCGAATTTAC
CATTACTTGGAATCTATTTAAAACCTCTTTAATCTTATGATA

[0001297] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1471. This T2 control element has the DNA sequence

[0001298] Seq. Id. = 190 Position = 1 to 95

[0001299] CAACTAACAACCGTATCGAATTTACCATTACTTGGAAATCTATTTAAAACCT
CTTTAATCTTGTGATAATAAATTCTAATCGATTCGTGACTTAT

[0001300] This long T1/T2 double stranded DNA loop modulates
the expression of the following genes

MJ1402	MJ1403	MJ1404	MJ1405	MJ1406
MJ1407	MJ1408	MJ1409	MJ1410	MJ1411
MJ1412	MJ1413	MJ1414	MJ1415	MJ1416
MJ1417	MJ1418	MJ1419	MJ1420	

[0001301] This long T1/T2 double stranded DNA loop modulates
the expression of the following C1/C2 short loops

[0001302] A C1/C2 short loop on chromosome 1 whose identifier
is 1448 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop is expressed as
a RNA single strand that is 3'UTR to the gene MJ1401 and has
the DNA sequence

[0001303] Seq. Id. = 191 Position = 1 to 122

[0001304] TTATAGAACATTATGAAGCTTTTTACTCAACTAACAACCGTATCGAATTTAC
CATTACTTGGAAATCTATTTAAAACCTCTTTAATCTTATGATAATAAATTCTAATCGATTCTG
TGAATTAT

[0001305] A C1/C2 short loop on chromosome 1 whose identifier
is 1470 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop is expressed as
a RNA single strand that is 3'UTR to the gene MJ1420 and has
the DNA sequence

[0001306] Seq. Id. = 192 Position = 1 to 116

[0001307] TTATAGAACATTATGAAGCTTTTTACTCAACTAACAACCGTATCGAATTTAC
CATTACTTGGAATCTATTTAAAACCTCTTTAATCTTGTGATAATAAATTCTAATCGATTCTG
TG

[0001308] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001309] A C1/C2 short loop on chromosome 1 whose identifier is 1470 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene MJ1420 and has the DNA sequence

[0001310] Seq. Id. = 193 Position = 1 to 116

[0001311] TTATAGAACATTATGAAGCTTTTTACTCAACTAACAACCGTATCGAATTTAC
CATTACTTGGAATCTATTTAAAACCTCTTTAATCTTGTGATAATAAATTCTAATCGATTCTG
TG

[0001312] The match between the T1 sequence and the C1/C2 sequence is

[0001313] Seq. Id. = 193 Position = 1 to 89

[0001314] TTATAGAACATTATGAAGCTTTTTACTCAACTAACAACCGTATCGAATTTAC
CATTACTTGGAATCTATTTAAAACCTCTTTAATCTT

[0001315] The match between the T2 sequence and the C1/C2 sequence is

[0001316] Seq. Id. = 193 Position = 28 to 116

[0001317] CAACTAACAACCGTATCGAATTTACCATTACTTGGAATCTATTTAAAACCT
CTTTAATCTTGTGATAATAAATTCTAATCGATTCTGTG

[0001318] A C1/C2 short loop on chromosome 1 whose identifier
is 1448 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene MJ1401 and has the DNA
sequence

[0001319] Seq. Id. = 194 Position = 1 to 122

[0001320] TTATAGAACATTATGAAGCTTTTACTCAACTAACAACCGTATCGAATTTAC
CATTACTTGGAATCTATTTAAAACCTCTTTAATCTTATGATAATAAATTCTAATCGATTCTG
TGACTTAT

[0001321] The match between the T1 sequence and the C1/C2
sequence is

[0001322] Seq. Id. = 194 Position = 1 to 95

[0001323] TTATAGAACATTATGAAGCTTTTACTCAACTAACAACCGTATCGAATTTAC
CATTACTTGGAATCTATTTAAAACCTCTTTAATCTTATGATA

[0001324] The match between the T2 sequence and the C1/C2
sequence is

[0001325] Seq. Id. = 194 Position = 29 to 99

[0001326] CAACTAACAACCGTATCGAATTTACCATTACTTGGAATCTATTTAAAACCT
CTTTAATCTT

[0001333] Seq. Id. = 196 Position = 1 to 77

[0001334] AATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTCG
AGGAGAACTTCTAGTATATTCTGTA

[0001335] This long T1/T2 double stranded DNA loop modulates
the expression of the following genes

YBL005W-B	TS(AGA)B	YBL004W	YBL003C	YBL002W
YBL001C	YBR001C	YBR002C	YBR003W	YBR004C
YBR005W	YBR006W	YBR007C	YBR008C	YBR009C
YBR010W	YBR011C	YBR012C		

[0001336] This long T1/T2 double stranded DNA loop modulates
the expression of the following C1/C2 short loops

[0001337] A C1/C2 short loop on chromosome 2 whose identifier
is 298 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop is expressed as
a RNA single strand that is 3'UTR to the gene YBL005W-B and
has the DNA sequence

[0001338] Seq. Id. = 5197 Position = 1 to 342

[0001339] ATCTATTACATTATGGGTGGTATGTTGGAATAAAAATCCACTATCGTCTATC
AACTAATAGTTATATTATCAATATATTATCATATACGGTGTTAAGATGATGACATAAGTTAT
GAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAA
TGAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCA
TTTGTAGGATTCCTATATCCTTGAGGAGAACTTCTAGTATATTCTGTATACCTAATATTATA
GCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTC

[0001340] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[0001341] A C1/C2 short loop on chromosome 2 whose identifier is 298 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YBL005W-B and has the DNA sequence

[0001342] Seq. Id. = 198 Position = 1 to 342

[0001343] ATCTATTACATTATGGGTGGTATGTTGGAATAAAAAATCCACTATCGTCTATC
AACTAATAGTTATATTATCAATATATTATCATATACGGTGTTAAGATGATGACATAAGTTAT
GAGAAGCTGTCATCGAAGTTAGAGGAAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAA
TGAAACATATAAAACGGAATGAGGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCA
TTTTGAGGATTCCTATATCCTTGAGGAGAACTTCTAGTATATTCTGTATACCTAATATTATA
GCCTTTATCAACAATGGAATCCCAACAATTATCTCAACATTC

[0001344] The match between the T1 sequence and the C1/C2 sequence is

[0001345] Seq. Id. = 198 Position = 23 to 276

[0001346] TGTGGAATAAAAAATCCACTATCGTCTATCAACTAATAGTTATATTATCAAT
ATATTATCATATACGGTGTTAAGATGATGACATAAGTTATGAGAAGCTGTCATCGAAGTTAG
AGGAAGCTGAAGTGCAAGGATTGATAATGTAATAGGATAATGAAACATATAAAACGGAATGA
GGAATAATCGTAATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCTT
GAGGAGAACTTCTAGT

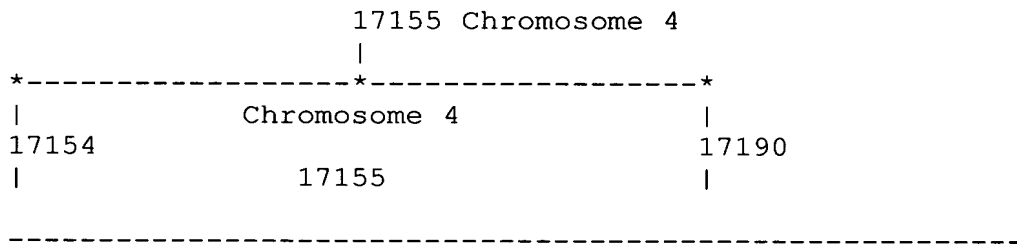
[0001347] The match between the T2 sequence and the C1/C2 sequence is

[0001348] Seq. Id. = 198 Position = 210 to 259

[0001349] AATATTAGTATGTAGAAATATAGATTCCATTTTGAGGATTCCTATATCCT

[0001350] Example of a multi-celled self-limiting connectron -
C. elegans

[0001351] In this example the existence of the T1-T2 (293-320) long loop is controlled by C1/C2 (298) short loop. The expression of this C1/C2 short loop is controlled by the existence of the T1-T2 (293-320) long loop. The existence of this T1-T2 long loop is itself determined by the expression of the C1/C2 (298) short loop. The C1/C2 (298) short loop is the self-limiting connectron.



[0001352] A double stranded DNA loop of length 89.919 kilobases on chromosome 4 is bounded on the left by a T1 sequence whose identifier is 17154. This T1 control element has the DNA sequence

[0001353] Seq. Id. = 199 Position = 1 to 29

[0001354] AAATTTCCGGCAAATCGGCAAACCTGGCAA

[0001355] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 17190. This T2 control element has the DNA sequence

[0001356] Seq. Id. = 200 Position = 1 to 29

[0001357] AATTTGCCGATTGCGGAATTTGTCGACA

[0001358] This long T1/T2 double stranded DNA loop modulates the expression of the following genes

R08C7.11	M01H9.2	M01H9.3	M01H9.4	M01H9.1
ZK180.1	ZK180.2	ZK180.3	ZK180.4	ZK180.5
ZK180.6	ZK185.3	ZK185.2		

[0001359] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001360] A C1/C2 short loop on chromosome 4 whose identifier is 17155 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene R08C7.1 and has the DNA sequence

[0001361] Seq. Id. = 201 Position = 1 to 56

[0001362] AAATTTCCGGCAAATCGGCAAAC TGGCAATTTGCCGATTTGCCGAATTTGTC
GACA

[0001363] A C1/C2 short loop on chromosome 4 whose identifier is 17171 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene ZK180.2 and has the DNA sequence

[0001364] Seq. Id. = 202 Position = 1 to 56

[0001365] TGGAAATTTTCAGAATTTCAATTTTAATCGGCAAATTTGTACGCATCCTATGA
ATTT

[0001366] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001367] A C1/C2 short loop on chromosome 4 whose identifier is 17155 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene R08C7.1 and has the DNA sequence

[0001368] Seq. Id. = 203 Position = 1 to 56

[0001369] AAATTTCCGGCAAATCGGCAAACCTGGCAATTTGCCGATTTGCCGAATTTGTC
GACA

[0001370] The match between the T1 sequence and the C1/C2 sequence is

[0001371] Seq. Id. = 203 Position = 1 to 29

[0001372] AAATTTCCGGCAAATCGGCAAACCTGGCAA

[0001373] The match between the T2 sequence and the C1/C2 sequence is

[0001374] Seq. Id. = 203 Position = 28 to 56

[0001375] AATTTGCCGATTTGCCGAATTTGTGCGACA

9. Geneless connectrons exist in single-celled and multi-celled eukaryotes

[0001376] Normally T1-T2 long loops contain genes whose expression is regulated by the existence of the long loop. When a T1-T2 long loop does not contain any genes it is described as being "geneless". The existence of the T1-T2 long loop is itself controlled by one or more C1/C2 short loops that may be on the same or different chromosomes. The geneless T1-T2 long loops must contain one or more C1/C2 short loops.

[0001377] Example of a single-celled geneless connectron - *S. cerevisiae*

[0001378] In this example the existence of the T1-T2 (1537-1559) long loop is controlled by three C1/C2 (3789, 5289 and 5753) short loops. The expression of 21 C1/C2 (1538 through 1558) short loops are controlled by the existence of the T1-T2 (1537-1559) long loop.

```

3789 Chromosome 9
5289 Chromosome 12
5753 Chromosome 13
|
*-----*-----*
|           Chromosome 4           |
1537                                           1559
|           1538 through 1558           |
|                                           |
-----
```

[0001379] A double stranded DNA loop of length 4.825 kilobases on chromosome 4 is bounded on the left by a T1 sequence whose identifier is 1537. This T1 control element has the DNA sequence

[0001380] Seq. Id. = 204 Position = 1 to 362

[0001381] ATGAGATATATGTGGGTAATTAGATAAATTGTTGGGATTCCATTGTTGATAAA
GGCTATAATATTAGGTATACAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATA
AAAGGGAATCTGCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAAT
ATTCATTGATCCTATTACATTATCAATCCTTGCGTTTTAGCTTCCACTAATTTAGATGACTA
TTTCTCATCATTTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATAC
TAGTTAGTAGATGATAGTTGATTTTTATTCCAACATACCACCCATAATGTAATAGATCTAAT

[0001382] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1559. This T2 control element has the DNA sequence

[0001383] Seq. Id. = 205 Position = 1 to 362

[0001384] ATGAGATATATGTGGGTAATTAGATAAATTGTTGGGATTCCATTGTTGATAAA
GGCTATAATATTAGGTATACAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATA
AAAGGGAATCTGCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAAT
ATTCATTGATCCTATTACATTATCAATCCTTGCGTTTTAGCTTCCACTAATTTAGATGACTA
TTTCTCATCATTTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATAC
TAGTTAGTAGATGATAGTTGATTTTTATTCCAACATACCACCCATAATGTAATAGATCTAAT

[0001385] There are no genes controlled by this T1/T2 loop.

[0001386] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001387] A C1/C2 short loop on chromosome 4 whose identifier is 1538 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001388] Seq. Id. = 206 Position = 1 to 387

[0001389] ATGAGATATATGTGGGTAATTAGATAATTGTTGGGATTCCATTGTTGATAAA
GGCTATAATATTAGGTATACAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATA
AAAGGGAATCTGCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAAT
ATTCATTGATCCTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTA
TTTCTCATCATTTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATAC
TAGTTAGTAGATGATAGTTGATTTTTATTCCAACATACCACCCATAATGTAATAGATCTAAT
GAATCCATTTGTTTGTTAATAGTTT

[0001390] This T1-T2 loop also modulates the C1/C2 short loops
numbered 1539 to 1557

[0001391] A C1/C2 short loop on chromosome 4 whose identifier
is 1558 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop has the DNA
sequence

[0001392] Seq. Id. = 207 Position = 1 to 307

[0001393] AGCTTCTCATAACTTATGTCATCATCTTAACACCGTATATGATAATATATTG
ATAATATAACTTGTTGGAATAAAAAATCAACTATCATCTACTAACTAGTATTTACGTTACTAG
TATATTATCATATACGGTGTTAGAAGATGACGCAAATGATGAGAAATAGTCATCTAAATTAG
TGGAAGCTGA...GTCTATCTGGCGAATATAAATTTTTACGCTACACACGTCATCGACATCT
AAATATGACAGTCGCTGAACTGTTCTTAGATATCCATGCTATTTATGAAGAACAACAGGGAT
CGAGAAACAG

[0001394] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[0001395] A C1/C2 short loop on chromosome 9 whose identifier
is 3789 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YIL059C and has the DNA
sequence

[0001396] Seq. Id. = 208 Position = 1 to 176

[0001397] TTTATATGTTAATATTCATTGATCCTATTACATTATCAATCCTTGCGTTTCA
GCTTCCACTAATTTAGATGACTATTTCTCATCATTTGCGTCATCTTCTAACACCGTATATGA
TAATATACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATTCCAACAGTAT

[0001398] The match between the T1 sequence and the C1/C2
sequence is

[0001399] Seq. Id. = 208 Position = 1 to 172

[0001400] TTTATATGTTAATATTCATTGATCCTATTACATTATCAATCCTTGCGTTTCA
GCTTCCACTAATTTAGATGACTATTTCTCATCATTTGCGTCATCTTCTAACACCGTATATGA
TAATATACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATTCCAACA

[0001401] The match between the T2 sequence and the C1/C2
sequence is

[0001402] Seq. Id. = 208 Position = 1 to 172

[0001403] TTTATATGTTAATATTCATTGATCCTATTACATTATCAATCCTTGCGTTTCA
GCTTCCACTAATTTAGATGACTATTTCTCATCATTTGCGTCATCTTCTAACACCGTATATGA
TAATATACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATTCCAACA

[0001404] A C1/C2 short loop on chromosome 12 whose identifier
is 5289 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YLR301W and has the DNA
sequence

[0001405] Seq. Id. =209 Position = 1 to 325

[0001406] GGTGAATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATAT
TAGGTATGTAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCT
GCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAATATTCATTGATC
CTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCAT
TTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAGA
TGATAGTTGATTTTTATTCCAACAC

[0001407] The match between the T1 sequence and the C1/C2
sequence is

[0001408] Seq. Id. =209 Position = 62 to 317

[0001409] AGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATC
TGCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAATATTCATTGAT
CCTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCA
TTTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAG
ATGATAGTTGATTTTTATTCCAACA

[0001410] The match between the T2 sequence and the C1/C2
sequence is

[0001411] Seq. Id. =209 Position = 86 to 324

[0001412] AGGATTTAGGAATCCATAAAAGGGAATCTGCAATTCTACACAATTCTATAAA
TATTATTATCATCGTTTTATATGTTAATATTCATTGATCCTATTACATTATCAATCCTTGCG
TTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCATTTGCGTCATCTTCTAACACCGTA
TATGATAATATACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATTCCAAC
A

[0001413] A C1/C2 short loop on chromosome 13 whose identifier
is 5753 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YMR044W and has the DNA
sequence

[0001414] Seq. Id. = 210 Position = 1 to 334

[0001415] TTGAGAAATGGGGGAATGTTGAGATAATTGTTGGGATTCCATTGTTGATAAA
GGCTATAATATTAGGTATACAGAATATACTAGAAGTTCTCCTCAAGGATATAGGAATCCTCA
AAATGGAATCTATATTTCTACATACTAATATTACGATTATTCCTCATTCCGTTTTATATGTT
TCATTATCCTATTACATTATCAATCCTTGCACTTCAGCTTCCTCTAACTTCGATGACAGCTT
CTCATAACTTATGTCATCATCTTAACACCGTATATGATAATATATTGATAATATAACTATTA
GTTGATAGACGATAGTGGATTTTTATTCCAACAT

[0001416] The match between the T1 sequence and the C1/C2
sequence is

[0001417] Seq. Id. = 210 Position = 22 to 95

[0001418] AGATAATTGTTGGGATTCCATTGTTGATAAAGGCTATAATATTAGGTATACA
GAATATACTAGAAGTTCTCCTC

[0001419] The match between the T2 sequence and the C1/C2
sequence is

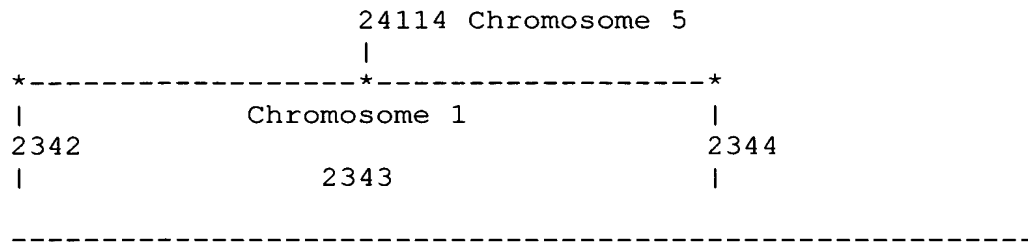
[0001420] Seq. Id. = 210 Position = 28 to 101

[0001421] TTGTTGGGATTCCATTGTTGATAAAGGCTATAATATTAGGTATACAGAATAT
ACTAGAAGTTCTCCTCAAGGAT

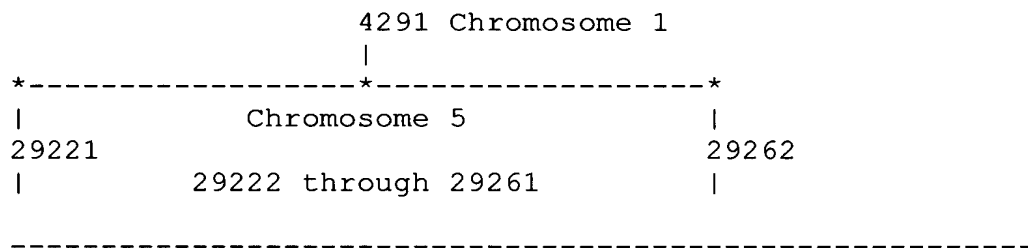
[0001422] Two examples of multi-celled geneless connectrons -
C. elegans

[0001423] In the first example the existence of the T1-T2
(2342-2344) long loop is controlled by the C1/C2 (24114) short
loop. The expression of one C1/C2 (2343) short loop is

controlled by the existence of the T1-T2 (2342-2344) long loop.



[0001424] In the second example the existence of the T1-T2 (29221-29262) long loop is controlled by the C1/C2 (24114) short loop. The expression of one C1/C2 (2343) short loop is controlled by the existence of the T1-T2 (2342-2344) long loop.



[0001425] A double stranded DNA loop of length 67.059 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 2342. This T1 control element has the DNA sequence

[0001426] Seq. Id. = 211 Position = 1 to 37

[0001427] TGAAAACTACAGTAATTCTTTAAATGACTACTGTAGC

[0001428] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 2344. This T2 control element has the DNA sequence

[0001429] Seq. Id. = 212 Position = 1 to 37

[0001430] CTACTGTAGCGCTTGTGTCGATTTACGGGCTCGATTT

[0001431] There are no genes controlled by this T1/T2 loop.

[0001432] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001433] A C1/C2 short loop on chromosome 1 whose identifier is 2343 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001434] Seq. Id. = 213 Position = 1 to 61

[0001435] TCGACACAAGCGCTACAGTAGCTATTTAAAGAATTACTGTAGTTTTCGCTAC
GAGATATTT

[0001436] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001437] A C1/C2 short loop on chromosome 5 whose identifier is 24114 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene C13F10.5 and has the DNA sequence

[0001438] Seq. Id. = 214 Position = 1 to 68

[0001439] GCGAAACTACAGTAATTCTTTAAATGACTACTGTAGCGCTTGTGTCGATT
ACGGGCTCGATTTTCG

[0001440] The match between the T1 sequence and the C1/C2
sequence is

[0001441] Seq. Id. = 214 Position = 3 to 38

[0001442] GAAACTACAGTAATTCTTTAAATGACTACTGTAGC

[0001443] The match between the T2 sequence and the C1/C2
sequence is

[0001444] Seq. Id. = 214 Position = 29 to 65

[0001445] CTACTGTAGCGCTTGTGTCGATTACGGGCTCGATTT

[0001446] A double stranded DNA loop of length 41.297 kilo-
bases on chromosome 5 is bounded on the left by a T1 sequence
whose identifier is 29221. This T1 control element has the
DNA sequence

[0001447] Seq. Id. = 215 Position = 1 to 62

[0001448] TTTAAATTTCCCGCCAAAAATTGACTGAAAATTTGGATTTTCTTTCCAAAA
TTGACAGAAA

[0001449] This double stranded DNA loop is bounded on the
right by a T2 control element whose identifier is 29262. This
T2 control element has the DNA sequence

[0001450] Seq. Id. = 216 Position = 1 to 31

[0001451] TGAAAATTTGAATTTCCCGCCAAAAATTAAC

[0001452] There are no genes controlled by this T1/T2 loop.

[0001453] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001454] A C1/C2 short loop on chromosome 5 whose identifier is 29222 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001455] Seq. Id. = 217 Position = 1 to 58

[0001456] AATTTCCCGCCAAAAATTGACTGAAAATTTGGATTTTCTTTCCAAAAATTGACAGAAA

[0001457] This T1-T2 loop also modulates the C1/C2 short loops numbered 29223 to 29260

[0001458] A C1/C2 short loop on chromosome 5 whose identifier is 29261 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001459] Seq. Id. = 218 Position = 1 to 54

[0001460] AAAATTGACTGAAAATTTGAATTTCCAGCCAAAAATTGACTGAAAATTTGAATT

[0001461] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001462] A C1/C2 short loop on chromosome 1 whose identifier is 4291 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene Y43F8C.5 and has the DNA sequence

[0001463] Seq. Id. = 219 Position = 1 to 317

[0001464] AAAATTAAGTGAATTTGAATTTCCCGCCAAAAATTGACTGAAAATTTGAA
TTTCCCGCCAAAAAAATTGACTGAAAATTTGAATTTCCCGCCAAAAATTGACTGAAAATTT
GAATTTCCCGCCAAAAATTAATTGAAAATTTGAATTTCCCGCCAAAAATTAATTGAACTTT
GAATTTTCAA...ATTTCCCGCCAAAAATTAATTGAACTTTGAATTTTCAAATTTCCCGCC
AAAAATTGACTGAAAATTTGAATTTCCCGCCAAAAATTAATTGAAAATTTGAATTTTGAAT
TTCCCGCCAAAAATGACTGA

[0001465] The match between the T1 sequence and the C1/C2 sequence is

[0001466] Seq. Id. = 219 Position = 229 to 260

[0001467] AAATTTCCCGCCAAAAATTGACTGAAAATTTG

[0001468] The match between the T2 sequence and the C1/C2 sequence is

[0001469] Seq. Id. = 219 Position = 63 to 104

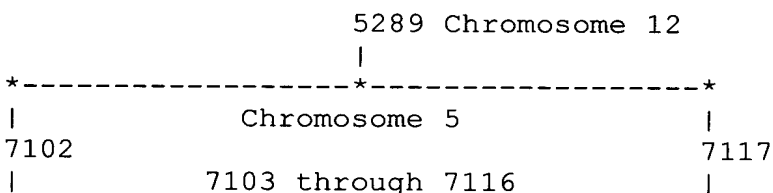
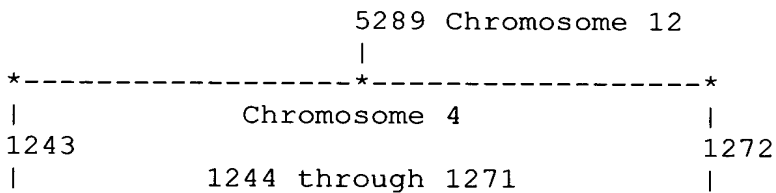
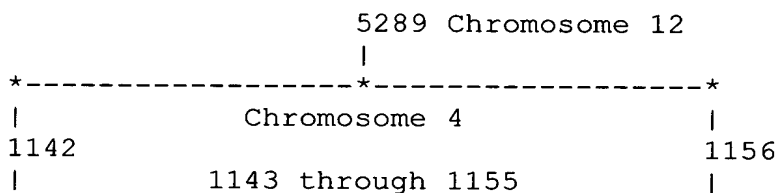
[0001470] AAAAAATTGACTGAAAATTTGAATTTCCCGCCAAAAATTGA

10. One connectron controls many geneless connectrons in single-celled and multi-celled eukaryotes

[0001471] One C1/C2 short loop can control the existence of many geneless T1-T2 long loops.

[0001472] Example of a single-celled geneless connectron - *S. cerevisiae*

[0001473] In this example the existence of the three T1-T2 (1142-1156, 1242-1272 and 7102-7117) long loops is controlled by the C1/C2 (5289) short loop.



[0001474] A double stranded DNA loop of length 5.337 kilobases on chromosome 4 is bounded on the left by a T1 sequence whose identifier is 1142. This T1 control element has the DNA sequence

[0001475] Seq. Id. = 220 Position = 1 to 318

[0001476] ATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATATTAGGT
ATGTAGATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCTGCAATT
CTACACAATTCTATAAATATTATTATCATCATTTTATATGTTAATATTCATTGATCCTATTA
CATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCATTTGCGT
CATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAGATGATAG
TTGATTTTTATTCCAACA

[0001477] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1156. This T2 control element has the DNA sequence

[0001478] Seq. Id. = 221 Position = 1 to 295

[0001479] TTTTAATAAGGCAATAATATTAGGTATGTAGATATACTAGAAGTTCTCCTCC
AGGATTTAGGAATCCATAAAAGGGAATCTGCAATTCTACACAATTCTATAAATATTATTATC
ATCATTTTATATGTTAATATTCATTGATCCTATTACATTATCAATCCTTGCGTTTCAGCTTC
CACTAATTTAGATGACTATTTCTCATCATTTGCGTCATCTTCTAACACCGTATATGATAATA
TACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATTCCAACAAGAA

[0001480] There are no genes controlled by this T1/T2 loop.

[0001481] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001482] A C1/C2 short loop on chromosome 4 whose identifier is 1143 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001483] Seq. Id. = 222 Position = 1 to 349

[0001484] ATTTTGAGATAATTGTTGGGATTCCATTTTTAATAAGGCAATAATATTAGGT
ATGTAGATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCTGCAATT
CTACACAATTCTATAAAATATTATTATCATCATTTTTATATGTTAATATTCATTGATCCTATTA
CATTATCAAT...CTCTAAGTCTCATTGCCTTTGTGCCAAAAAATCTGTTTCTAAATTTCTC
TTCATTTGTAGACTTAATTATACTGATCGTTGATCTACTATCAGTAAGTAAGCCTTTAATAA
TTGGTTTCTTGTTAAGTTCTTGCACAAGGTGACTGAGGTTATTCAATAGCGG

[0001485] This T1-T2 loop also modulates the C1/C2 short loops
numbered 1144 to 1154

[0001486] A C1/C2 short loop on chromosome 4 whose identifier
is 1155 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop has the DNA
sequence

[0001487] Seq. Id. = 223 Position = 1 to 69

[0001488] GAGGAGAACTTCTAGTATATCTACATACCTAATATTATTGCCTTATTAAAAA
TGGAATCCCAACAATTA

[0001489] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[0001490] A C1/C2 short loop on chromosome 12 whose identifier
is 5289 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YLR301W and has the DNA
sequence

[0001491] Seq. Id. = 224 Position = 1 to 324

[0001492] GGTGAATTTTGAGATAATTGTTGGGATTCCATTTTTAATAAGGCAATAATAT
TAGGTATGTAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCT
GCAATTCTACACAATTCTATAAAATATTATTATCATCGTTTTATATGTTAATATTCATTGATC

CTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCAT
TTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTACGTAAATACTAGTTAGTAGAT
GATAGTTGATTTTTATTCCAACAC

[0001493] The match between the T1 sequence and the C1/C2
sequence is

[0001494] Seq. Id. = 224 Position = 6 to 64

[0001495] ATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATATTAGGT
ATGTAGA

[0001496] The match between the T2 sequence and the C1/C2
sequence is

[0001497] Seq. Id. = 224 Position = 33 to 64

[0001498] TTTTAATAAGGCAATAATATTAGGTATGTAGA

[0001499] A double stranded DNA loop of length 5.251 kilo-
bases on chromosome 4 is bounded on the left by a T1 sequence
whose identifier is 1243. This T1 control element has the DNA
sequence

[0001500] Seq. Id. = 225 Position = 1 to 366

[0001501] CGTGTTTTATCTCATGTTGTTTCGTTTTGTTATTGAGATATATGTGGGTAATT
AGATAATTGTTGGGATTCCATTGTTGATAAAGGCTATAATATTAGGTATACAGAATATACTA
GAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCTGCAATTCTACACAATTCTAT
AAATATTATTATCATCGTTTTATATGTTAATATTCATTGATCCTATTACATTATCAATCCTT
GCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCATTTGCGTCATCTTCTAACACC

GTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATTCC
AACA

[0001502] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 1272. This T2 control element has the DNA sequence

[0001503] Seq. Id. = 226 Position = 1 to 273

[0001504] TGAGATATATGTGGGTAATTAGATAATTGTTGGGATTCCATTGTTGATAAAG
GCTATAATATTAGGTATACAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAA
AAGGGAATCTGCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAATA
TTCATTGATC...TATACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATT
CCAACAGTTATAAGGTTGTTTCATATGTGTTTTATGAA

[0001505] There are no genes controlled by this T1/T2 loop.

[0001506] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001507] A C1/C2 short loop on chromosome 4 whose identifier is 1244 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001508] Seq. Id. = 227 Position = 1 to 327

[0001509] TTTATCTCATGTTGTTTCGTTTTGTTATTGAGATATATGTGGGTAATTAGATA
ATTGTTGGGATTCCATTGTTGATAAAGGCTATAATATTAGGTATACAGAATATACTAGAAGT
TCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCTGCAATTCTACACAATTCTATAAATA
TTATTATCAT...GTCTCGATGTAGTATACGTATAAATTATTACCTGATACTTCATCTCTAA
GTCTCATTGCCTTTGTGCCAAAAAATCTGTTTCTAAATTCTCTTCATTTGTAGACTTAATT
ATACTGATCGTTGATCTACTATCAGTAAGT

[0001510] This T1-T2 loop also modulates the C1/C2 short loops numbered 1245 to 1270

[0001511] A C1/C2 short loop on chromosome 4 whose identifier is 1271 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001512] Seq. Id. = 228 Position = 1 to 309

[0001513] TGTTGTATCTCAAAATGAGATATGTCAGTATGACAATACGTCATCCTAAACG
TTCATAAAACACATATGAAACAACCTTATAACTGTTGGAATAAAAATCAACTATCATCTACT
AACTAGTATTTACGTTACTAGTATATTATCATATACGGTGTTAGAAGATGACGCAAATGATG
AGAAATAGTC...CAACAATGGAATCCCAACAATTATCTAATTACCCACATATATCTCATGG
TAGCGCCTGTGCTTCGGTTACTTCTAAGGAAGTCCACACAAATCAAGATCCGTTAGACGTTT
CAGCTTCCAAAA

[0001514] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001515] A C1/C2 short loop on chromosome 12 whose identifier is 5289 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene YLR301W and has the DNA sequence

[0001516] Seq. Id. = 229 Position = 1 to 325

[0001517] GGTGAATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATAT
TAGGTATGTAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCT
GCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAATATTCATTGATC
CTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCAT
TTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAGA
TGATAGTTGATTTTTATTCCAACAC

[0001518] The match between the T1 sequence and the C1/C2 sequence is

[0001519] Seq. Id. = 229 Position = 62 to 317

[0001520] AGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATC
TGCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAATATTCATTGAT
CCTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCA
TTTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAG
ATGATAGTTGATTTTTATTCCAACA

[0001521] The match between the T2 sequence and the C1/C2 sequence is

[0001522] Seq. Id. = 229 Position = 62 to 317

[0001523] AGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATC
TGCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAATATTCATTGAT
CCTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCA
TTTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAG
ATGATAGTTGATTTTTATTCCAACA

[0001524] A double stranded DNA loop of length 5.296 kilobases on chromosome 15 is bounded on the left

[0001525] by a T1 sequence whose identifier is 7102. This T1 control element has the DNA sequence

[0001526] Seq. Id. = 230 Position = 1 to 365

[0001527] CATGATTAATATGACCAATCGGCGTGTGTTTTGAAAAGTGGGTGAATTTG
AGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATATTAGGTATGTAGAATGTACTAG

AAGTTCTCCTCAAGGATTTAGGAATCCATGAAAGGGAATCTGCAATTCTACACAATTCTATA
AATATTATTATCATCATTTTATATGTTAATATTCATTGATCCTATTACATTATCAATCCTTG
CGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCATTTGCGTCATCTTCTAACACCG
TATATGATAATATACTAGTAACGTAAATACTAGTTAGTAGATGATAGTTGATTTTTATTCCA
ACA

[0001528] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 7117. This T2 control element has the DNA sequence

[0001529] Seq. Id. = 231 Position = 1 to 365

[0001530] TGAAAAGTGGGTGAATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGG
CAATAATATTAGGTATGTAGAATGTACTAGAAGTTCTCCTCAAGGATTTAGGAATCCATGAA
AGGGAATCTGCAATTCTACACAATTCTATAAATATTATTATCATCATTTTATATGTTAATAT
TCATTGATCCTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATT
TCTCATCATTTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTA
GTTAGTAGATGATAGTTGATTTTTATTCCAACAGTTTATATACCTCTCTTATTTAGTATAA
GAA

[0001531] There are no genes controlled by this T1/T2 loop.

[0001532] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001533] A C1/C2 short loop on chromosome 15 whose identifier is 7103 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001534] Seq. Id. = 232 Position = 1 to 357

[0001535] AAGAACATTGCTGATGTGATGACAAAACCTCTCCGATAAAAACATTTAAAC
TATTAAC TAACAAATGGATTCATTAGATCTATTACATTATGGGTGGTATGTTGGAATAAAAA

TCAACTATCATCTACTAACTAGTATTTACGTTACTAGTATATTATCATATACGGTGTTAGAA
GATGACGCAAATGATGAGAAATAGTCATCTAAATTAGTGGAAGCTGAAACGCAAGGATTGAT
AATGTAATAGGATCAATGAATATTAACATATAAAATGATGATAATAATATTTATAGAATTGT
GTAGAATTGCAGATTCCCTTTCATGGATTCCCTAAATCCTTGAGGAGAACTTCTAGTA

[0001536] This T1-T2 loop also modulates the C1/C2 short loops
numbered 7104 to 7115

[0001537] A C1/C2 short loop on chromosome 15 whose identifier
is 7116 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop has the DNA
sequence

[0001538] Seq. Id. = 233 Position = 1 to 66

[0001539] CCATTCTGTGGAGGTGGTACTGAAGCAGGTTGAGGAGAGACATGATGATGGT
TCTCTGGAACAGCT

[0001540] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[0001541] A C1/C2 short loop on chromosome 12 whose identifier
is 5289 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene YLR301W and has the DNA
sequence

[0001542] Seq. Id. = 234 Position = 1 to 325

[0001543] GGTGAATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATAT
TAGGTATGTAGAATATACTAGAAGTTCTCCTCGAGGATTTAGGAATCCATAAAAGGGAATCT
GCAATTCTACACAATTCTATAAATATTATTATCATCGTTTTATATGTTAATATTCATTGATC
CTATTACATTATCAATCCTTGCGTTTCAGCTTCCACTAATTTAGATGACTATTTCTCATCAT

TTGCGTCATCTTCTAACACCGTATATGATAATATACTAGTAACGTAAATACTAGTTAGTAGA
TGATAGTTGATTTTTATTCCAACAC

[0001544] The match between the T1 sequence and the C1/C2
sequence is

[0001545] Seq. Id. = 234 Position = 1 to 66

[0001546] GGTGAATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATAT
TAGGTATGTAGAAT

[0001547] The match between the T2 sequence and the C1/C2
sequence is

[0001548] Seq. Id. = 234 Position = 1 to 66

[0001549] GGTGAATTTTGAGATAATTGTTGGGATTCCATTTTAAATAAGGCAATAATAT
TAGGTATGTAGAAT

[0001550] Example of a multi-celled geneless connectron - C.
elegans

[0001551] In this example the existence of the three T1-T2
(1142-1156, 14840-15042 and 15365-15627) long loops is
controlled by the C1/C2 (16760) short loop.

```

                        16760 Chromosome 4
                        |
*-----*-----*
|           Chromosome 4           |
1142                                     1156
|           3103 through 3119       |

```

```

              16760 Chromosome 4
              |
*-----*-----*
|           Chromosome 4           |
14840                                     15042
|           14841 through 15041     |

```

```

              16760 Chromosome 4
              |
*-----*-----*
|           Chromosome 5           |
15365                                     15627
|           15366 through 15626     |

```

[0001552] A double stranded DNA loop of length 15.894 kilobases on chromosome 1 is bounded on the left by a T1 sequence whose identifier is 3101. This T1 control element has the DNA sequence

[0001553] Seq. Id. = 235 Position = 1 to 33

[0001554] CAAATCGGCAAATTGCCGGAATTGAACATTTCC

[0001555] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 3120. This T2 control element has the DNA sequence

[0001556] Seq. Id. = 236 Position = 1 to 54

[0001557] AAACGATTTTTCGGCAAATCGGCAAATTGCCGGAATTGTAATTTCCGGCAA
AT

[0001558] There are no genes controlled by this T1/T2 loop.

[0001559] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001560] A C1/C2 short loop on chromosome 1 whose identifier is 3103 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001561] Seq. Id. = 237 Position = 1 to 55

[0001562] TTAAAATTTCCGGCAAATCGGCAAATTGGCAGAAATGAAACTCACGGCAAAT
CGG

[0001563] This T1-T2 loop also modulates the C1/C2 short loops numbered 3104 to 3118

[0001564] A C1/C2 short loop on chromosome 1 whose identifier is 3119 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001565] Seq. Id. = 238 Position = 1 to 61

[0001566] CCCGCATTTTTTTGTAGATCAAACCGTAATGGGACGGCCTGGCAACACGTGAT
TTTCCAAAT

[0001567] The expression of genes in this T1/T2 long loop is controlled by the following C1/C2 short loops.

[0001568] A C1/C2 short loop on chromosome 4 whose identifier is 16760 controls the expression of the genes in this T1/T2 long loop. This C1/C2 short loop is expressed as a RNA single strand that is 3'UTR to the gene T23E1.2 and has the DNA sequence

[0001569] Seq. Id. = 239 Position = 1 to 124

[0001570] GGCAAATTGCCGAAATTGAACATTTCCGGCAAATCGGCAAATTGCCGGAATT
GAACATTTCCGGCAAATCGGCAAATTGCCGGAATTGAACATTTCCGGCAAATCGGCAAATTG
CCGGAATTGA

[0001571] The match between the T1 sequence and the C1/C2
sequence is

[0001572] Seq. Id. = 239 Position = 30 to 62

[0001573] CAAATCGGCAAATTGCCGGAATTGAACATTTCC

[0001574] The match between the T2 sequence and the C1/C2
sequence is

[0001575] Seq. Id. = 239 Position = 23 to 53

[0001576] TTTCCGGCAAATCGGCAAATTGCCGGAATTG

[0001577] A double stranded DNA loop of length 86.977 kilo-
bases on chromosome 3 is bounded on the left by a T1 sequence
whose identifier is 14840. This T1 control element has the
DNA sequence

[0001578] Seq. Id. = 240 Position = 1 to 141

[0001579] AAAAATTTCCGGCAAGTCGGCAATTTTCCGAAAATGAAAATTTCCGGCAAAT
CGGCAAATTGCCGGAATTGAAAATTCCTGGCAAATCAGCAAATTTGCGGCAAATCGGCAATT
TGCCGAAAATGAAAATTTCCGGCAAAT

[0001580] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 15042. This T2 control element has the DNA sequence

[0001581] Seq. Id. = 241 Position = 1 to 98

[0001582] CAAATCGGTAGGTAAATTGGCCAAACTTGAAAATTTCCGGCAAATCGGCAAA
TTCCGCGAACTGAACATTTCCGGCAAATCGGCAAATTGCTCGAACT

[0001583] There are no genes controlled by this T1/T2 loop.

[0001584] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001585] A C1/C2 short loop on chromosome 3 whose identifier is 14841 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001586] Seq. Id. = 242 Position = 1 to 141

[0001587] AAAAATTTCCGGCAAGTCGGCAATTTTCCGAAAATGAAAATTTCCGGCAAAT
CGGCAAATTGCCGGAATTGAAAATTCCTGGCAAATCAGCAAATTTGCGGCAAATCGGCAATT
TGCCGAAAATGAAAATTTCCGGCAAAT

[0001588] This T1-T2 loop also modulates the C1/C2 short loops numbered 14842 to 15040

[0001589] A C1/C2 short loop on chromosome 3 whose identifier is 15041 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001590] Seq. Id. = 243 Position = 1 to 55

[0001591] CGGCAATTGCCGTTCCGGCAATTTGCCAATTTGCCGGAAATTTTCAATTCCGG
CAA

[0001592] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[0001593] A C1/C2 short loop on chromosome 4 whose identifier
is 16760 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene T23E1.2 and has the DNA
sequence

[0001594] Seq. Id. = 244 Position = 1 to 124

[0001595] GGCAAATTGCCGAAATTGAACATTTCCGGCAAATCGGCAAATTGCCGGAATT
GAACATTTCCGGCAAATCGGCAAATTGCCGGAATTGAACATTTCCGGCAAATCGGCAAATTG
CCGGAATTGA

[0001596] The match between the T1 sequence and the C1/C2
sequence is

[0001597] Seq. Id. = 244 Position = 22 to 55

[0001598] ATTTCCGGCAAATCGGCAAATTGCCGGAATTGAA

[0001599] The match between the T2 sequence and the C1/C2
sequence is

[0001600] Seq. Id. = 244 Position = 17 to 45

[0001601] TGAACATTTCCGGCAAATCGGCAAATTGC

[0001602] A double stranded DNA loop of length 98.488 kilobases on chromosome 3 is bounded on the left by a T1 sequence whose identifier is 15365. This T1 control element has the DNA sequence

[0001603] Seq. Id. = 245 Position = 1 to 336

[0001604] AAAATTTCCGGCAAATCGGCAATTTGCCAAAAATTGAAATTTCCGGCAAATC
GGCAATTTGTCAAAAATGAAAATTTCCGGCAAATCGGCAAATTGCCGAAAATGAAAATTTCC
GGCAAATCGGCAAACCTCCGGAACCTGAAAATTTCCGGCAAATCGGCAATTTGCCATAAATGA
ACATTTCCGG...GGCGAAAATTAAAATTTCCGCCATATCGGCAATTTGCCAAAAAATTAAA
ATTTCCGGCAAATCGGCAAATTGCCGGAATTCAAAATTTCCGGCAAACCGGCAAATTGCCGG
AACTCAAATTTCCCGGCAAATCAGCAAATTGCCGGAATT

[0001605] This double stranded DNA loop is bounded on the right by a T2 control element whose identifier is 15627. This T2 control element has the DNA sequence

[0001606] Seq. Id. = 246 Position = 1 to 68

[0001607] TGGCAAACCGGCAAATTGCCGGAATTGAACATTTCCGGCAAATCGGCAATTT
GCCGGAATTGAAATTT

[0001608] There are no genes controlled by this T1/T2 loop.

[0001609] This long T1/T2 double stranded DNA loop modulates the expression of the following C1/C2 short loops

[0001610] A C1/C2 short loop on chromosome 3 whose identifier is 15366 controls the expression of the genes of one or more other T1/T2 long loops. This C1/C2 short loop has the DNA sequence

[0001611] Seq. Id. = 247 Position = 1 to 60

[0001612] TGCCGATTGCGGAAATTTTCATTTTCGGCAATTTGCCGATTGCGGAAA
TTTTCATT

[0001613] This T1-T2 loop also modulates the C1/C2 short loops
numbered 15366 to 15624

[0001614] A C1/C2 short loop on chromosome 3 whose identifier
is 15625 controls the expression of the genes of one or more
other T1/T2 long loops. This C1/C2 short loop has the DNA
sequence

[0001615] Seq. Id. = 248 Position = 1 to 54

[0001616] TCAAGCAAATTGTCAAATTCGCGGAACATAACATTTCCGGCAAATCGGCAAA
TT

[0001617] The expression of genes in this T1/T2 long loop is
controlled by the following C1/C2 short loops.

[0001618] A C1/C2 short loop on chromosome 4 whose identifier
is 16760 controls the expression of the genes in this T1/T2
long loop. This C1/C2 short loop is expressed as a RNA single
strand that is 3'UTR to the gene T23E1.2 and has the DNA
sequence

[0001619] Seq. Id. = 249 Position = 1 to 124

[0001620] GGCAAATTGCCGAAATTGAACATTTCCGGCAAATCGGCAAATTGCCGGAATT
GAACATTTCCGGCAAATCGGCAAATTGCCGGAATTGAACATTTCCGGCAAATCGGCAAATTG
CCGGAATTGA

[0001621] The match between the T1 sequence and the C1/C2 sequence is

[0001622] Seq. Id. = 249 Position = 22 to 52

[0001623] ATTTCCGGCAAATCGGCAAATTGCCGGAATT

[0001624] The match between the T2 sequence and the C1/C2 sequence is

[0001625] Seq. Id. = 249 Position = 35 to 75

[0001626] CGGCAAATTGCCGGAATTGAACATTTCCGGCAAATCGGCAA

References

1. "Functional and structural units in the chromomere", Resnik, N. A.; Yampol, G. P.; Kiseleva, E. V.; Khristolyubova N. B. and Gruzdev, A. D., *Genetica* 83:293-299, 1991.
2. "A Silence that speaks volumes", Gura, T., *Nature* 404:804-808.
3. "RNA Interference", Sharp, P. A. and Zamore, P. D., *Science* 287:2431-2433.
4. "Genetic Requirements for Inheritance of RNAi in *C. elegans*", Grishok, A.; Tabara, H.; Mello, C. C., *Science* 287:2494-2496.
5. "An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells", Hammond, S. M.; Bernstein, E.; Beach, D. and Hannon, G. J., *Nature* 404:293-296.
6. "A genetic link between co-suppression and RNA Interference in *C. elegans*", Ketting, R. F. and Plasterk, R. H., *Nature* 404:296-298